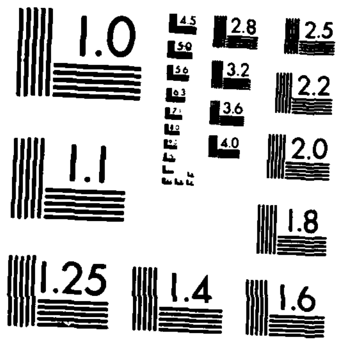
AD-A174 726 OPTIMAL COMBINATION OF INFORMATION FROM MULTIPLE SOURCES(U) MASSACHUSETTS INST OF TECH CAMBRIDGE DEPT OF MECHANICAL ENGINEERING M B MENDEL ET AL 31 JUL 86 1/1

UNCLASSIFIED N00014-83-K-0193 F/G 5/8 NL



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

12

Man-Machine
Systems Laboratory

Optimal Combination of Information from Multiple Sources

Max B. Mendel
Thomas B. Sheridan

July 1986

CONTRACT N00014-83-K-0193
WORK UNIT NR 196-179
ENGINEERING PSYCHOLOGY PROGRAMS
OFFICE OF NAVAL RESEARCH
ARLINGTON, VIRGINIA 22217

Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A174 726

REPORT DOCUMENTATION PAGE

1a REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3 DISTRIBUTION/AVAILABILITY OF REPORT: APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED
2b DECLASSIFICATION/DOWNGRADING SCHEDULE	
4 PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: MASSACHUSETTS INSTITUTE OF TECHNOLOGY; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: ENGINEERING PSYCHOLOGY PROGRAM OFFICE OF NAVAL RESEARCH (CODE 442EP)
6c. ADDRESS (City, State, and ZIP Code): 77 MASSACHUSETTS AVENUE CAMBRIDGE, MA 02139	7b ADDRESS (City, State, and ZIP Code): 800 North Quincy Street Arlington VA 22217
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: Office of Naval Research; 8b. OFFICE SYMBOL (If applicable)	9 PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: N00014-83-K-0193
8c. ADDRESS (City, State, and ZIP Code): 800 North Quincy St. Arlington, VA 22217	10 SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO; PROJECT NO NT; TASK NO 196; WORK UNIT ACCESSION NO 179

11 TITLE (Include Security Classification)

OPTIMAL COMBINATION OF INFORMATION FROM MULTIPLE SOURCES

12 PERSONAL AUTHOR(S)

Max B. Mendel, Thomas B. Sheridan

13a TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL (PART 3)	FROM Mar 83 TO July 86	1986 July 31	86

16 SUPPLEMENTARY NOTATION

17 COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Supervisory control; mental models; decision aids; human-computer interaction

19 ABSTRACT (Continue on reverse if necessary and identify by block number)

A computer decision aiding system for debiasing and combining information from multiple sources (e.g. human experts, sensors) is proposed. The algorithm is based on six assumptions that apply when the sources are relatively knowledgeable with respect to the operator on the variable of interest, and the operator is willing to base his evaluation of their performance on a previously selected (finite) sequence of so called calibration variables. It is also assumed that the operator is interested in maximizing gains, that is, he wishes to act in an optimal or Bayesian manner. An experiment with two human sources of information was conducted to evaluate the performance of the aiding system under a variety of loss funtions. On a family of bilinear loss functions, the output of the aid was found to perform better than a naive scheme like simply believing the information the two sources gave. The combination rule was also found to perform better than the output to any individual source.

20 DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS	21 ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a NAME OF RESPONSIBLE INDIVIDUAL: Thomas B. Sheridan	22b TELEPHONE (Include Area Code): 617-253-2228; 22c OFFICE SYMBOL: 3-346

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete

SECURITY CLASSIFICATION OF THIS PAGE

MAN-MACHINE SYSTEMS LABORATORY
Department of Mechanical Engineering
Massachusetts Institute of Technology
Cambridge MA 02139

Optimal Combination of Information from Multiple Sources.

Max B. Mendel
Thomas B. Sheridan

ABSTRACT

A computer decision aiding system for debiasing and combining information from multiple sources is proposed. The algorithm is based on six assumptions that apply when the sources are relatively knowledgeable with respect to the operator on the variable of interest, and the operator is willing to base his evaluation of their performance on a previously selected (finite) sequence of so called calibration variables. It is also assumed that the operator is interested in maximizing gains, that is, he wishes to act in an optimal or Bayesian manner. An experiment with two human sources of information was conducted to evaluate the performance of the aiding system under a variety of loss funtion. On a family of bilinear loss functions, the output of the aid was found to perform better functions than a naive scheme like simply believing the information the two sources gave. The combination rule was also found to perform better than the output to any individual source.

1. INTRODUCTION

The essence of the presence of an operator in a supervisory control situation is that he can choose and execute a control action. According to the principles of decision theory this preference the operator has among control action can be conveniently described in terms of what he *knows* about his system and what he *wants* to achieve, the two being postulated to be independent of each other. What he knows, his beliefs, can be represented numerically in terms of a "subjective probability function" over the possible states of the system, and what he wants, his values, in terms of a (subjective) "utility function" or "loss function" over the possible consequences of his actions. Thus, the operator's behavior can be described in terms of two notions, which in turn can be uniquely represented by these two functions; in much the same way as, for example, the behavior of particles can be represented conveniently in terms of its mass and the forces working on which, in turn, can be uniquely represented by real numbers.

The companion report by Charny and Sheridan [1986] on a computer aiding technique for satisficing, investigates the aiding of the operator in establishing what he wants. This report shall be devoted to aiding the operator to gather knowledge about the state of the system, or, equivalently, to develop an accurate mental model of values of the state variables of the system.

The way we obtain knowledge is by observation. However, observing the state of a system in a typical supervisory control situation is far less straightforward than it is to check, say, whether the grass is still green, or, to take a favorite statistical paradigm, to check whether a coin has landed with heads or tails upward. Observations in supervisory control situations usually come to the operator in an indirect way via some source of information like the equipment on his control panel, "intelligent" devices like fault location systems, computer data bases or human experts. An important practical difference is that, if the information comes to the operator via an information source, it may be deformed or biased in some way, whereas in the previous, direct case -barring hallucinogenic states of mind- what you see is what you get. For example, a source may systematically overestimate certain quantities, or it might be systematically overconfident. Therefore, since an information source may not give an accurate reflection of the state of the system, the operator needs to know its "characteristics" to make optimal use of its information.

If the operator has more than one source at his disposal, the situation becomes even more involved. Sources of information, like any set of observations, are mutually redundant to some degree. Just as we know that a double check of the color of grass will not add much to our beliefs, the operator might know that one source typically duplicates another source, making the one irrelevant in light of the other.

Therefore, in order to incorporate the information from all the available sources within his knowledge base, the operator needs to know something about the characteristics of the sources individually, and as a whole. With this he can debias and weigh the information from each source to combine them, together with his own opinions, into one reliable and consistent mental model of the system.

2. PROBLEM STATEMENT

The purpose of this paper is to propose a computer aiding system that can, in some well described situations, take over this task of combining information from different sources. We will develop the mathematical model for the general case, that is for any number of sources. We will also show the results of the computer implementation, that is the actual aiding system, when applied to an experimental setup with two human sources of information. The experiment is a tentative one, designed to illustrate the concepts and workings of the theory and the aiding system. Full scale experiments simulating actual industrial situations are presently being considered and will be reported on at a later stage. For the experimental results, we then show that under a large class of loss functions -i.e. no matter what he wants- the operator is better off (saves money) believing the output of the aid than simply accepting the advice of the sources.

The way an operator would actually combine information from different sources, or, more appropriately, how he should do it, is a very intricate affair involving many factors. Of course we cannot hope to automate this procedure in general. What we can (and should) do is to state a number of plausible and operational assumptions under which an optimal solution to this problem exists. If the operator agrees that the assumptions accurately describe his particular situation, he can let the computer aiding system take over the task.

For the aid to be useful, the assumptions have to apply to frequently occurring and interesting situations. As will be made precise in the course of this manuscript, the aiding system based on these assumptions pertains to situations in which the information sources are knowledgeable with respect to the operator himself on the states on which he seeks advice. As information normally costs something (money or trouble or any kind of loss), this situation is fairly common; if the operator thinks he knows more than his information sources, he need not bother with them at all.

All in all we will put forward six assumptions. To cast the problem into a numerical framework it is necessary to make three "form" assumptions. The first form assumption, F1, will impose some necessary structure on the state-space and the operator. The second form assumption, F2, fixes a standard way of describing the information flow from the source to the operator. Form assumption three, F3, then introduces a numerical way of measuring the characteristics of the sources, or the process of calibration. We then proceed to carry through the necessary inference steps. To accomplish this we introduce three "inference assumptions". I1 fixes a starting point, that is a (prior) set of beliefs when the operator has as yet no data pertinent to the characteristics of the sources. I2 determines the relation between the calibration measurements. I3, finally, fixes how sensitive the starting point in I1 is to these measurements. These assumptions are sufficient to guarantee the existence of a optimal solution, i.e. a (posterior) set of beliefs over the states of the system given the information from the information sources and their characteristics as evidenced by the calibration data.

The paper is organized as follows: Chapter 3 introduces the three form assumption and chapter 4 the three inference assumptions. Chapters 5 and 6 present the solution to the combination algorithm that results from these assumptions and presents the results when the model is applied to the experimental setup. Chapter 8 contains the evaluation of the model, i.e. the decision aid, and chapter 8 rounds everything off with a summary and conclusions. The setup and results of the experiment are contained in appendices 1 to 3. Appendix 4 contains some derivations that are left out in the main text in order not to distract the attention from the main argument. Appendix 5 contains a list of symbols used in the course of the document.

3. ANALYSIS AND NOTATION.

3.1 F1: Form of the system and the operator

At a particular moment in time, the system can be viewed to be in any of a number of states, in other words, anywhere in state-space. In order for the remaining assumptions to make sense, and for plain convenience we shall impose some structure on this space. The only structure that the other assumptions require is a topological one, namely that the state-space is a simply ordered set.[1] This assumption is rather minimal, and I do not think any practical system of interest exists for which this condition does not hold. In fact, the following, much stronger condition holds so frequently that we shall assume that there exists a, for all practical purposes, real-valued random variable defined on the state-space, which we call $\tilde{x}_0$, on the outcome, x_0, of which the operator bases his decisions.[2] This situation occurs so often that I believe that the loss in generality is amply counterbalanced by the gain in familiarity and ease of expression.

The experiment involved fifty-three (real-valued) variables taken out of the Guinness book of world records and selected for their pertinence to mechanical engineering (see appendix 1). The maximum design sway of the Eiffel tower, given its length (question number 47) is an example of a random variable defined on the space of extreme states of the Eiffel tower. Similarly, the deviation from the parallel of the towers of the Humber Estuary bridge (question 48) is a variable defined on the states of that bridge, and so for the capacity of the aqueduct of Carthage (question 49) etc. etc.

If we plan to aid the operator, we will have to assume that there are some basic principles of logic the operator wants to adhere to. To this end, we will assume that the operator wishes to be a Bayesian; specifically, that he wishes to comply to Savage's [1954] axioms. These axioms consist of the minimal equipment to ensure consistency, or coherency is it is more often referred to, of the operator's behavior. As mentioned previously, they also allow us to describe the operator's behavior in terms of (subjective) probabilities and losses, where the operator will wish to choose the action that carries the least expected loss. (This part of F1 justifies that reasoning in the introduction.) It is in this sense that the decision aid will be optimal. Thus F1 ensure optimality of the model with respect to any loss function. This does not mean that the aid cannot be wrong. What is does mean is that there is no systematic way in which it can lose independent of the actual state of the system. It can be shown (see de Finetti [1974]) that any other set of axioms can be used as a "money pump", i.e. there exist a systematic way in which it can lose independent of the actual state of the system. For the supervisory control situation these properties are obviously extremely desirable.

Since the first part of F1 allows us to parametrize the state-space by $\tilde{x}_0$, and the second part allows us to summarize the relevant aspects of the operator's mental model of the system by probabilities, we can write the operator mental model as a distribution over $\tilde{x}_0$ given all his knowledge (except the information from the sources and the data about their characteristics). We comply with tradition by assuming the conditioning on his knowledge to be implicit and simply write $F(x_0)$. This distribution is called the a priori predictive distribution. Since, in this

1. A relation $\leq$ is called a simple ordering (or a linear order, a weak order or an order relation) on a set S if (and only if) for all x, y, and z in S
   1. (comparability) Either $x \leq y$, or $y \leq x$.
   2. (transitivity) If $x \leq y$, and $y \leq z$, then $x \leq z$.
2. We distinguish between random variables and their outcomes by providing the former with a tilde.

document, no confusion is possible we shall simply call it the operator's prior (on x_0).[3]

Now suppose that the operator obtains information from k different sources, which we label $A, B, \ldots, K$.[4] Suppose also that the operator has some idea about the reliability of sources $A, \ldots, K$. His mental model of the system can now be no longer represented by $F(\tilde{x}_0)$ but should be represented by his distribution over $\tilde{x}_0$ *given* the information *and* the characteristics of its sources, or,

$$F(x_0 \mid \textit{Information from } A, \ldots, K, \textit{ Characteristics of } A, \ldots, K) \tag{3.1}$$

Let us call this distribution, the predictive distribution a posteriori, the posterior (on $\tilde{x}_0$).

Thus it will be our goal to evaluate (3.1). To accomplish this we still need to perform some analysis. Indeed, the vague terms *Information from* $A, \ldots, K$ and *Characteristics of* $A, \ldots, K$ need to be replaced by precise numerical concepts. This will be covered by F2 and F3, respectively.

Although (3.1) will give the operator a complete representation of the relevant beliefs he need have, it is not at all obvious that he is able to use (3.1) effectively. We will not discuss this delicate issue here, but refer the reader to the companion report by Roseborough and Sheridan [1986] where this is investigated at large, both theoretically and experimentally.

3.2 F2: Form of the information

Ideally, each source of information should give its entire probability distribution over the variable of interest, $\tilde{x}_0$. Obviously, and this especially so if the sources are humans, this is not feasible. Instead we will assume that each source is capable, possibly in a noisy and/or systematically biased way, to give some points, say m of them, of its distribution. The continuous case should then be viewed as a limiting case as $m \to \infty$. For our purposes it will prove to be most convenient to let these points be a source's fractiles of its distribution.[5]

Thus, suppose we can elicit m numbers from each source which the source calls its fractiles to probability values that have been preset. To keep track of these numbers we shall order them in increasing magnitude and label them with the lower case of the source in question. In other words, we define for source A a (row) vector, or ordered set $\mathbf{a} = (a_1, \ldots, a_m)$ that contains A's m fractiles on the variable $\tilde{x}_0$. Similarly for source B to K.

If there is an m for which the above scheme works, we can replace *Information from* $A, \ldots, K$ in (3.1) by the ordered sets $\mathbf{a}, \ldots, \mathbf{k}$ that contain the sources' fractiles. In other words, expression (3.1) can now be written as

$$F(x_0 \mid \mathbf{a}, \ldots, \mathbf{k}, \textit{ Reliability of } A, \ldots, K). \tag{3.2}$$

The next question is a practical one: how many and which fractiles can one reasonably ask of an information source? Of course the answer will depend on the nature of the source. Most

3. The notation $F(\cdot)$ is used for any (cumulative) distribution, i.e., any marginal distribution of the distribution over all states of the system. Densities will be represented by $f(\cdot)$ and the measure itself by $p(\cdot)$.

4. More precisely, it is the states of information of the sources that are labeled such.

5. An α-fractile of a distribution, $F(x_0)$ is a value of $\tilde{x}_0$ such that the probability that the true value is below that value equals α, or, simply $F^{-1}(\alpha)$.

equipment gives a single reading that represents its mean or its median and the distribution around this number is provided by the manufacturer. In this case an infinite number of fractiles are available. For human experts or expert systems the horn is a little less plenty. Often human experts give their advice in terms of a single estimate which we can consider to be his median. Just one fractile seems a bit meager, but if k is not too small -if there are enough experts- this kind of information will still be useable. A more reasonable approach is to ask the expert to give an estimate, and let that be the median of his distribution, and also an 80% confidence or credible interval around that estimate. In fractile words, the source gives what it claims are its .1-, .5- and .9-fractiles. Here we have of course that $a=(a_1, a_2, a_3)$, a_1 being the .1-fractile and so on. Similarly source B's information is embedded in $b = (b_1, b_2, b_3)$. The situation is portrayed in figure 3.1 for some distribution that source A might have.

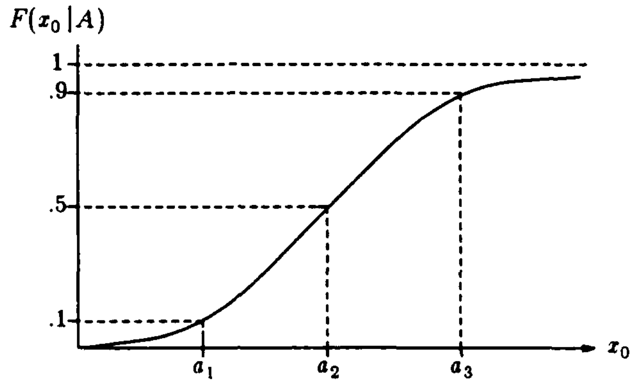


Figure 3.1. Hypothetical fractiles $a = (a_1, a_2, a_3)$ of source A for the values set in the experiment.

This was the method that was actually employed in the experiment. Two human subjects, source A and source B, had to provide their .1-, 5- and .9-fractiles on the fifty-three random variables in appendix 1. The experiment was designed to simulate the situation in which experts advise an operator as closely as could be expected within the confines of this tentative experiment. The subjects, both Ph.D. candidates in Mechanical Engineering at MIT, were therefore allowed to use all standard engineering equipment like desk calculators. Of course they were not allowed to simply look up the true value.

3.3 F3: Characteristics and calibration

Of course it is one thing asking for someone's fractiles, and another thing to actually get them. Even assuming that the source can introspect sufficiently to give its fractiles, there may still be reason to doubt the unbiasedness of its opinions, or to which degree it is redundant in the light of another available information source. We will, in this chapter, propose a way to measure the characteristics of the sources no matter what their cause is. To justify some of the choices made in this chapter, we need anticipate some of the results of the following chapters, hoping that the reader will be willing to temporarily accept some statements as reasonable.

The way to determine the characteristics of a source of information is to compare its opinions about the value of a variable to the true value of that variable. A little, but not much more

precisely said, we want to measure the "distance" between the opinions and the true value. Of course, after one variable we cannot say much; the source may have been particularly lucky or unlucky. However, after a large number of variables to which we compare the advice and the true values, it is not unreasonable to expect to converge to some belief about the characteristics of the source, especially if these variables are chosen from some area of expertise of the source.

We will first propose a way of measuring "distances" between the fractiles of a source on a variable and its true value. Suppose then that, at the time the operator obtains the advice on $\tilde{x}_0$, he has n variables to he not only has the sources' advice, but to which he also has the true values. We shall call this kind of variables calibration variables. To distinguish these from $\tilde{x}_0$, the decision variable at that moment whose true value he naturally does not possess, we shall denote them by $\tilde{x}_1, \ldots, \tilde{x}_n$. In practice these variables can either be selected beforehand such that the operator knows their true values, maybe by putting the sources on a simulator of the system, or they could have been previous decision variables to which the true value has meanwhile become known; if a source predicts the Dow Jones index for tomorrow, than tomorrow this value will be known. In the latter case the system gradually learns the characteristics of the sources. Naturally, combinations of these approaches can also be considered.

Concerning the division of the experimental data into decision and calibration variables, we can imagine that each of the fifty-three variables can be viewed as a decision variable, with the preceding variables being used as calibration variables, i.e. at variable number $n+1$ we assume that the values to the previous n are meanwhile known. This is borne out in the legends to the graphs in appendix 3. At the first variable, there are no calibration variables; if $\tilde{x}_0$ is the first variable then $n=0$. In general, when $\tilde{x}_0$ is the $n+1^{th}$ variable, then the operator has the previous n variables as calibration variables at his disposal. Thus the experiment can be viewed as a sequential operation where at each subsequent variable the operator and his aiding system have one more calibration variable at their disposal to determine the characteristics of the sources more precisely.

Now their are of course many ways of defining the distance between the advice on a variable and its true value. From form considerations alone, however, the following can be said. Consider thereto the variables in the experiment. One variable is measured in miles per hour, another in degrees Kelvin etc. If we wish to make any kind of comparison between these distances, we have to introduce some way of measuring distance that is independent of the particular scale the variable itself is measured in. One way to do this is to simply note between which fractiles the true value falls. No matter how the scale is changed, this statistic will be invariant. If we obtain three fractiles from each source, then the true value can fall in any of four interfractile ranges, or bins as we shall call them: the true value is either smaller than a_1, or between a_1 and a_2, or between a_2 and a_3, or larger than a_3. We shall call the first interfractile range the zeroth bin, the second the first bin and so on to the third bin.

For notational purposes it is convenient to introduce, for every variable $\tilde{x}_i$ ($i=0,1, \ldots, n$) a new random variable $\tilde{z}_i$ that indicates in which bin the true value has fallen or might fall. The way in which this is actually done is chosen purely for mathematical convenience and need not concern us here (see appendix 4). We shall call $\tilde{z}_i$ the calibration indicator for variable $\tilde{x}_i$. The importance of the calibration indicators, the $\tilde{z}_i$, is that they have the same scale for each i, although the scales of the corresponding $\tilde{x}_i$ may be totally incommensurable. This fact is essential for the inference as exposed in chapter 4.

When we have two sources, we have $((3+1)\cdot(3+1) =)$ sixteen different bins and the calibration indicators can be generalized accordingly. Let us agree to number the bins of the double-source case with two digit numbers, the first indicating in which (marginal) bin of source A the true value falls, and the second in which (marginal) bin of source B. We reserve the letter j to represent a generic bin number. For notational convenience, we shall think of j as a k-dimensional vector, $j=[j_a, \ldots, j_k]$, having an element for each source indicating in which

(marginal) bin of that source the true value fell. Thus, for the two-source case, if the true value falls in A's zeroth bin and in B's second bin, we will say it fell in bin $j=[j_a, j_b]=[0,2]$. Before we state F3 and proceed with the inference assumptions, we will first discuss intuitively why this way of measuring distance using calibration indicators is reasonable from an inferential point of view, that is, what we may hope to expect from it when using the inference assumptions to update the operator's distribution over $\tilde{x}_0$.

If a source were unbiased we would expect that, in the long run, 10% of the true values would fall in the zeroth and 10% in the third bin, since the source assigns probability .1 to these bins. Similarly 40% of the true values should fall in the first and 40% in the second bin. A source for which this is so is called well-calibrated, a rather unusual state of affairs. Suppose now that, for large n, most of the true values fall in the zeroth and third bin. Apparently, when the source says the events are unlikely (probability .1), they occur far too often, and when it says the events almost surely happen (probability .9: zeroth, first and second bin together) they do not happen often enough. Such a source can be called overconfident and we would, on the basis of this information, like to correct for this overconfidence when determining the posterior on $\tilde{x}_0$ by making his distribution proportionally "flatter". Similarly, a source can be underconfident by having too many true values fall in the middle two bins. If too many true values fall in the zeroth and first bin we would say that the source consistently underestimates the magnitude of the quantities, also something one could compensate for, etc. etc. Thus, in the single-source case the sequence of calibration indicators should be capable, under appropriate conditions, to identify a source's biases.

When the operator has two sources, A and B, at his disposal, the calibration indicators also measure, in a specific sense, the degree of interdependence between the two sources. It may be the case that, when A underestimates, B consistently overestimates. By measuring at the same time in which of A's and B's bin a true value falls, the sequence of calibration indicators can thus also indicate the interdependence between the two sources, that is, measure the characteristics of A and B as a group or panel of information sources.

Of course, the amount of different biases and interdependencies -i.e. characteristics- that can be identified is potentially infinite and we will, therefore, not attempt to do this. The important point is that, at least intuitively speaking, whatever the biases and interdependences may be, they will be partially embedded in the particular sequence of calibration indicators, $z_1, \ldots, z_n$, and this partiality will become more and more complete as n grows. These considerations suggest replacing *Characteristics of* $A, \ldots, K$ in (3.2) by a sequence of calibration indicators, $z_1, \ldots, z_n$. This sequence should enable the aid to correct for all the existing biases as much as possible for a given value of n. In the following we will indeed assume that we can do this, and thus write for (3.2)

$$F(x_0 \mid a, \ldots, k, z_1, \ldots, z_n). \tag{3.3}$$

This assumption entails that the operator's knowledge about the characteristics of the sources consists solely of the sequence of n calibration indicators. Often this will be approximately true, or at least that the operator is so insecure about his personal evaluations, that he is willing to disregard them for all but very small values of n. In other cases the operator will have to express his extra-calibration-indicator-sequence-knowledge in terms that are compatible with the calibration indicator sequence to incorporate this knowledge with the evidence from the calibration indicators. Some more will be said on this in appendix 4. Meanwhile, we will assume that from now on (3.3) represents the appropriate posterior.

It should be clear that calibration does not measure how much or little a source knows, but how well it knows what it knows, or better, how well it can state what it knows. If a source does not know much about a certain variable it should choose its fractiles far enough apart to reflect its ignorance. If it thinks it knows a lot, it should choose its fractiles close enough together. Too close will make the source overconfident, too far apart will make it underconfident. Whether

knowledgeable or ignorant, in general or on a particular variable, a source can be well-calibrated. An important underlying idea of the above reasoning, that will be made precise by I2, is that knowing what one knows is something that is relatively constant for each source and can be measured by determining the calibration indicators for many different variables, i.e., we expect our and the operator's beliefs concerning the characteristics of the sources to converge, as n grows, to the true biases of the sources.

4. INFERENCE

The previous section contained a lot of intuitive reasoning with many inherent assumptions like the existence and meaningfulness of hit-rates, convergence of the operator's beliefs and so on. Moreover, the arguments were based on the long run, i.e., for infinite n, which of course, strictly speaking, never applies in practice. Indeed, especially for human sources, calibration data are difficult to obtain, so solutions for the finite case must also be available. (Even if a large number of calibration indicators are available, the question would still remain which number is large enough.) We shall now state three assumptions which will provide the operator -if he finds them acceptable for his situation- with a unique posterior for any number n of calibration questions. The assumptions will also be seen to justify the intuitive reasoning of the previous section.

From the point of view of statistical inference, the posterior (3.3) can be analyzed in terms of two separate updatings: one on the information from the sources, and one on the characteristics of the sources. The actual inference will not be done on $\tilde{x}_0$, but on the corresponding calibration indicator, $\tilde{z}_0$. The corresponding distributions on $\tilde{x}_0$ will then be derived from these. Schematically we shall pass through the following inference steps

$$\begin{array}{ccccc} F(x_0) & \longrightarrow & F(z_0) & \longrightarrow & F(z_0 \,|\, z_1, \ldots, z_n) \\ & & \downarrow & & \downarrow \\ & & F(z_0 \,|\, a, \ldots, k) & \rightarrow & F(z_0 \,|\, a, \ldots, k, z_1, \ldots, z_n) \\ & & \downarrow & & \downarrow \\ & & F(x_0 \,|\, a, \ldots, k) & & F(x_0 \,|\, a, \ldots, k, z_1, \ldots, z_n) \end{array}$$

Figure 4.1. Schema of the steps involved in the inference to the posterior, $F(x_0 \,|\, a, \ldots, k, z_1, \ldots, z_n)$. The small arrow indicates an alternate route.

I1 will concern the form of $F(z_0)$ and how to obtain from it $F(z_0 \,|\, a, \ldots, k)$ and $F(x_0 \,|\, a, \ldots, k)$, the distribution over the decision indicator and variable, respectively, after the operator has obtained the information from the sources but has no information pertinent to their reliability. The remaining two assumptions will be concerned with the conditioning on $z_1, \ldots, z_n$ and the subsequent conditioning on $a, \ldots, k$.

4.1 I1: Prior beliefs

If one wants to get somewhere, one has to start off somewhere. Thus, if we want to calculate a posterior distribution, we first have to establish a prior distribution. The starting points in the above schema are $F(x_0)$ and $F(z_0)$, distributions which are neither conditioned on the information, nor on any outcomes of calibration indicators. For lack of better beliefs on the characteristics of the sources, owing to F3, it seems reasonable to give the sources the benefit of the doubt and to simply start off believing its statements. In other words, the operator assumes the sources are well-calibrated. The operator can then rely on the conditioning on the calibration indicators to disprove the source if necessary. If nothing else, the principle appeals to some notion of fairness, somewhat reminiscent of common practice in courtrooms where the accused is assumed innocent until proven otherwise by the evidence.

If the operator has one source at his disposal, it is clear what this assumption means, namely that $F(z_0)$ is simply equal to the probabilities to which the source gave his fractiles; in the experiment: .1, .5, .9 for the zeroth through the third bin. For the multi-source case, the interpretation of the

assumption has to be extended a little. Suppose, as in the experiment, that the operator has two sources at his disposal. As discussed under F3, this results in a total of sixteen bins which we can arrange in the following matrix

.1	00	.01	01	.04	02	.04	03	.01
.4	10	.04	11	.16	12	.16	13	.04
.4	20	.04	21	.16	22	.16	23	.04
.1	30	.01	31	.04	32	.04	33	.01
		.1		.4		.4		.1

Figure 4.2. Prior bin probabilities in the double-source four-fractile case, as in the experiment.

The bin numbers (the values of j) are shown in small print and the probabilities in normal print. The single-source interpretation fixes the marginal bin probabilities which are shown at the sides of the table. Of course these are not enough to specify all sixteen bin probabilities. We will take the statement of I1 to mean also that the operator considers the sources, a priori, to be "independent" in the sense that the probability of the true value falling in a certain bin of A and a certain bin of B is the product of the marginal bin probabilities. For example, the prior probability that the true value falls in the zeroth bin of both A and B is $.1 \cdot .1 = .01$. Extensions to the k-source case are obvious.

The side step is to condition $F(z_0)$ on the information from the sources to obtain $F(z_0 \mid a, \ldots, k)$. In the single-source case this is straightforward. Since $a, \ldots, k$ by F3 does not contain any information about the characteristics of the sources, it follows that, for instance, $F(z_0) = F(z_0 \mid a)$. Appendix 4 contains a more thorough discussion. In the multi-source case, however, things are slightly less straightforward. The problem is that, given the advice, not all bins remain possible under any reasonable $F(x_0)$. If, for instance, a_1 is smaller or equal than b_1 it cannot be, at the same time, larger than b_1. It follows that, given the advice, at most (3+3+1=) seven of the original sixteen bins remain (and at least four if the a_i all coincide with the b_i). The appropriate conditioning can, here be achieved by simply renormalizing the appropriate probabilities. This is expanded on in more detail in appendix 4.

The distribution $F(z_0 \mid a, \ldots, k)$ constrains the distribution $F(x_0 \mid a, \ldots, k)$ by prescribing the total amount of probability mass on the interfractile ranges or the bins. For instance, .4 of the probability should be assigned between a_1 and a_2. Their are a several ways to fix the shape of $F(x_0 \mid a, \ldots, k)$ on the bins itself. In the experiments we simply assumed that this distribution is uniform over the bins. In appendix 3-A these densities are shown by a dashed line for the case that only source A is available, $F(x_0 \mid a)$, and appendix 3-B shows the same when B is the only available source, $F(x_0 \mid b)$. Appendix 3-AB shows, again in dashed lines, $F(z_0 \mid a, b)$.

4.2 I2: order of the calibration indicators

In chapter 5 there were a several references to the relative frequency of true values falling in a certain bin or the hit rate of that bin. Let y_j be the number of true values falling in bin j. Then the these references imply that the ratios of the y_j to n are all that is needed of the entire sequence of calibration indicators $z_1, \ldots, z_n$ to determine a source's calibration. Statisticians would say that these ratios are considered to be sufficient statistics for the sequence of calibration indicators.

We will actually make a somewhat weaker assumption, namely that the counts y_j are sufficient

statistics for the sequence. (Clearly the ratios can be constructed from these counts. Thus if the counts are sufficient, the ratios are sufficient but not vice versa.) A logically equivalent, but intuitively more appealing, statement of this is that we -and hopefully the operator with us- consider the order in which the calibration indicators appear to be irrelevant.[6] In statistics, such variables are known as "exchangeable" random variables, since one may be exchanged for another (of course prior to obtaining the outcome) without changing the probabilities and, therefore, the decisions.

The assumption of exchangeability is a very typical assumption when one is measuring a quantity. When one flips a coin to determine the limiting frequency of say heads, when one makes multiple measurements of the same constant quantity and so on, one assumes that the order in which the outcomes will appear is irrelevant; a simple count of the total number of heads and tails is considered to be fully adequate.

As a direct notational result, we can now replace the sequence of calibration indicators, $z_1, \ldots, z_n$, by the counts of the hits in the different bins. For the single-source case with four different bins it is convenient to define the vector $y = [y_0, \ldots, y_3]$ (see also appendix 4). In the two-source case we can define y to be a $(m+1)\cdot(m+1)$-matrix, which, for m=3 is of the same dimensions as the matrix in 4.2. We will use the same subscripts for the elements of y. Thus y_{00} represents the number of true values that fell in A's and B's zeroth bin. In the general case, y becomes a k-dimensional array. The details are in appendix 4. For the present purposes it is sufficient to realize that the notation y represents that the counts of the true values falling in the $(m+1)^k$ bins. The posterior (3.3) can now be shortened with this notation as follows

$$F(x_0 \mid a, \ldots, k, y). \tag{4.1}$$

and similarly for all other distributions containing $z_1, \ldots, z_n$. Appendices 3-A and 3-B contain the two single-source cases. The counts that are available at the time that information on that particular decision variable is obtained are shown underneath the graphs.

Although the assumption of exchangeability seems innocuous enough, some surprisingly strong consequences exist. First of all, if one calibration indicator can be exchanged, a priori, for another, it follows that the operator's prior on the indicators must be equal. Mathematically, let z_i and z_l be any two calibration indicators, i,l=1,....,53, and pretend that the operator does not know their outcomes yet but assumes I2, then

$$p(z_i) = p(z_l). \tag{4.2}$$

I1 effectively takes care of this equality. It states not only that they are equal, but also to which distribution (the sources' claims). Using the same reasoning it also follows that the conditional probabilities of the z_i, the calibration indicators, are equal.[7] In particular, let $\tilde{z_i}$ and $\tilde{z_l}$ be any two calibration indicators. By exchangeability we thus have,

$$p(z_i \mid y) = p(z_l \mid y), \tag{4.3}$$

saying, in words, that the posteriors on calibration indicators are equal, given equal calibration data. Therefore, assuming exchangeability, the operator should not only consider the calibration indicators to be equally distributed a priori, but he should also change his beliefs on them in the same way given equal evidence. An incomplete and imprecise, but very suggestive way of saying the same is that the operator should be just as *sure* or *confident* about his beliefs in all the calibration indicators.

However, just as confident does not say how confident, like equally distributed does not say how

7. Actually, (4.2) is a special case of (4.3).

they are distributed. Therefore, we need an assumption that sets the level of confidence in the calibration indicators in the same spirit that I1 sets the distribution of the calibration indicators. In other words, we have to extend I1 to take care of all the probability dynamics. Given the structure imposed by I2 this is feasible and will be done by the following, and final assumption, I3.

4.3 I3: Confidence in calibration indicators

Since it is usually easier to identify extremes, let us begin by investigating the two extreme cases in which the operator is totally confident and the least confident about his prior on the calibration indicators.

If the operator is totally sure about the prior as postulated in I1, he will not be willing to change his beliefs no matter what calibration data are available. Mathematically we can say that

$$p(z_i \mid y) = p(z_i) \tag{4.4}$$

for all y. In other words, such an operator is unwilling to learn from experience since he considers he knows at all anyway a priori. Calibration data are completely useless for him and so would consequently be this research.

However, when an operator seeks advice he is typically not at all sure about his beliefs on the variable. The other extreme, i.e., wherein the operator is as unconfident as Bayesianly possible about his beliefs, would therefore be much more practical. An operator could reason that he is willing to accept the prior in I1, but that he will want to change it as quickly as possible on incoming evidence. This kind of assumption has the appeal of giving the data a maximum of influence on the posterior (and the prior a minimum).

Of course there are many shades of gray between the two extremes cited here. To establish which level of confidence in the prior on the calibration indicators the operator really wants will require input from him. We will assume here that the operator is indeed as unconfident as possible in the prior. The formal statement of this assumption relies upon de Finetti's theorem and is presented in appendix 4, which also contains a discussion on how to proceed when the operator is not willing to accept I3.

5. POSTERIOR ON THE CALIBRATION INDICATOR

The three form assumptions and the three inference assumption stated above are sufficient to determine a unique posterior distribution over the bins of $\tilde{x}_0$, i.e., over $\tilde{z}_0$. For the details of the derivation, the interested reader is referred to appendix 4. Here we shall merely list an approximate result for the case wherein $p(z_0) \leq .5$, which is the case in the experiment. Where z_0 denotes the true value of $\tilde{x}_0$ falling in bin j, we find that

$$p(z_0 \mid y) = \frac{y_j}{n + \frac{1}{p(z_0)}} \tag{5.1}$$

Conditioning on the information, $a, \ldots, k$ to pass to $F(z_0 \mid a,\ldots,k,\ y)$, can be achieved in exactly the way as was described under I1, when the objective was to pass from $F(z_0)$ to $F(z_0 \mid a, \ldots, k)$, by throwing out the impossible bins and renormalizing. See appendix 4 for the details.

Alternatively, we could have taken $F(z_0 \mid a, \ldots, k,\ y)$, as a starting point, as follows

$$p(z_0 \mid a, \ldots, k,\ y) = \frac{y_j + 1}{n + \frac{1}{p(z_0 \mid a, \ldots, k)}} \tag{5.2}$$

Thus, the posterior on a particular bin is a function of the prior, the count of hits in the bin and the total numbers of counts in the existing bins for that indicator. The prior is determined by I1 and the claims of the sources. The y_j's and n are determined by the outcomes of the previous calibration indicators together with I2. They are listed underneath the graphs in appendices 2 and 3. Thus, the posterior on the calibration indicator for the decision variable is uniquely determined by (5.1) or (5.2).

It is interesting to note that in (5.1), as n grows large and as long as $p(z_0) \neq 0$, the posterior on the bins approaches the ratio y_j/n, as conjectured in chapter 3.3. Thus, in the long run, the operator's posterior beliefs will indeed converge to the hit rates of the bins. For the finite case (i.e. all practical cases), however, the posterior is determined by both the prior and the calibration data, and the hit rates alone are insufficient. The solution extends the intuition in that the solution to the finite cases ($n < \infty$) is also given.

Appendix 2 shows the posterior on the calibration indicator of several decidion variables. On the horizontal axis the priors, $F(z_0 \mid a)$, $F(z_0 \mid b)$ or $F(z_0 \mid a,b)$, on the calibration indicators are plotted. On the vertical axis the posterior on the calibration indicators is shown. If the source is well calibrated, these should be equal for all variables; the points should all lie on the line making a 45 degree angle with both axes. Of course, on the first variable this is, by assumption, the case since there are no calibration data yet to disprove the source. Thus variables A-1, B-1 and AB-1 in appendix 2 all display this phenomenon. The values are indicated by little boxes.

At variable A-2 we see that the true value of the first variable fell in A's zeroth bin (as can also be verified in appendix 3A-1) so the posterior on the calibration indicator of A-2 is adjusted to make the zeroth bin somewhat more probable and the remaining bins proportionally less probable. At variable A-5, some more evidence has arrived that A overestimates (true values falling in zeroth bin) and the plot already takes on the typical shape of an overconfident source (too many hits falling in the extreme bins). At A-10 the source is just overconfident and at A-20 it looks as an overconfident underestimator. At A-53, the last variable, it is (provisionally) concluded that source A is somewhat overconfident and tend to underestimate the quantities in the experiment. The way the posterior says this is by lending more probability to the zeroth and third bin and, of these two, the most to the third bin.

One can see that initially the posterior on the calibration indicator is very sensitive to incoming calibration data. This is due to I3 which postulates that the operator is as unconfident as possible within the constraints of I2 about his prior beliefs as in I1. As he accumulates calibration data, however, the operator becomes more and more convinced of the characteristics of his source so at, say, variable A-40 his beliefs are significantly less sensitive to incoming calibration data. A previously performed Monte Carlo simulation shows that the posterior on the calibration indicators converges to its limit for all practical purposes after about 500 trials.

Also shown in appendix 2 are the calibration results for source B. It so happened that source B also had the true value of the first variable fall in its zeroth bin, so plot B-1 is equal to A-1. After fifty-three variables B shows significantly more overconfidence than A and also a strong tendency to underestimate. Both these characteristics are somewhat expected for our human sources. Overconfidence is a general bias that people have. Since the questions mostly involved records, one can conjecture that these tend to be unexpectedly high in general so that people tend to underestimate records. Comparing B-53 to A-53, we can say that source A is better calibrated than B. However, this does not necessarily imply that A is also more informative than B, either a priori or a posteriori, since being informative is also dependent on the location of the source's fractiles. This will be the subject of chapter 6.

Appendix 2AB contains the calibration plots for the two-source case. In general these contain seven data points. However, sometimes A and B each have a fractile at the same location, in which case there are only six distinct ones (e.g. on variable 10). Variable AB-1 is, of course, the familiar straight line under 45 degrees. The subsequent calibration plots show a similar kind of behavior as in the results of A and B alone.

For readers familiar with standard calibration plots it should be noted that there are two important differences between those and the plots in appendix 2. First of all, the results in appendix 2 are valid for all variables, i.e., no matter how large n, the number of available calibration indicators, is. Standard calibration plots are only valid for large n; strictly speaking only for $n=\infty$. Secondly the results for the two-source case are novel in that they are *not* obtained by simply adding the counts of the two sources for each bin. They are based on counts of all possible cross-bins and they measure calibration in the one-source sense plus the interdependence of the two sources. For these readers it may also be interesting to realize that the standard definition of calibration in the literature entails I2, that is, exchangeability of the calibration indicators. Indeed any definition that considers hit rates to determine, in the long run, calibration considers y to be a sufficient statistic which is equivalent to assuming exchangeability. The consequences of exchangeability are surprisingly strong though. Under I1 we discussed that the operator need not only consider his marginal priors on the indicators to be equal, but change them equally too. This is something that is insufficiently verified in the empirical psychological literature on calibration (see e.g. Lichtenstein et al. [1980] for a review to that date). As an extreme example, consider an information source that cheats in the following way. On $\tilde{x}_1$ it makes its zeroth bin ridiculously large by choosing a_1 to be extremely large. Obviously, everyone expects the true value to fall in the zeroth bin. On $\tilde{x}_2$ this source could stretch out another bin of its choosing. By dosing things right it could thus attain any calibration score it wants, including perfect calibration. The reason all arguments break down, though, is that I2 is violated. The operator's priors are not equal over the calibration indicators and thus they are not exchangeable. The violations of I2 can be more subtle. For instance, suppose the operator has the same prior on $\tilde{z}_0$ and the $\tilde{z}_1, \ldots, \tilde{z}_n$, and he is very unsure about these beliefs on the $\tilde{z}_1, \ldots, \tilde{z}_n$ but very sure on his beliefs on $\tilde{z}_0$ (e.g., this involves the outcome of some die which he has rolled a very large number of times). Now the calibration of the source on $z_1, \ldots, z_n$, especially if n is small, is not going to influence the operator opinions on $\tilde{z}_0$. This would have been completely different if $\tilde{z}_0$ was exchangeable with $\tilde{z}_1, \ldots, \tilde{z}_n$, i.e., if he was just as unconfident about his beliefs in z_0.

It would seem that strict exchangeability is unattainable in practice and that calibration of experts

and this decision aiding system is practically meaningless. We do not think that this is the right conclusion. The moral is, rather, that one has to be very careful in selecting a sequence $z_1, \ldots, z_n$ of calibration indicators and to try to approach the ideal as much as possible, or at least deviate from it in a controlled and informed way. To come back to the Newtonian analogy, the fact that an inertial reference frame is unattainable in practice does not make it a useless concept or often a practically viable assumption.

6. POSTERIOR ON THE DECISION VARIABLE

The calibration indicators are merely derived random variables designed to assist in the analysis. As stated in (3.1) through (3.3), our fundamental interest is in the posterior beliefs on the decision variable. In the experiment every variable has its turn as a decision variable, with the previous variables acting as calibration indicators.

The posterior on the calibration indicator of the decision variable in question constrains the possible posteriors on the variable since they prescribe the probability mass of each bin. However, they do not determine a unique posterior distribution on the decision variable. The way we obtain a posterior on the decision variable is to invoke F3 again. For the present purposes F3 implies that the operator wishes to change his opinions not any more than is implied by the calibration data. One way to achieve this mathematically is to maximize the entropy of the posterior with respect to the prior of the decision variable, under the constraints imposed by the posterior on the calibration indicator of that variable.

Since the exact shape of the prior is rather arbitrarily determined in I1 (uniform over the bins), the posterior will be just as arbitrary (but will not introduce extra arbitrariness). It can be shown that, in our case, the posterior is also uniform over the bins. Appendix 3 shows, in solid lines, the posterior distributions for the fifty-three variables. The posteriors to the information from source A alone are shown in appendix 3A. At variable A-1 the prior is equal to the posterior; the source, by lack of calibration data, has the benefit of the doubt. The true value is marked on the horizontal axis by a cross. Its value is listed in the legend of the graph, which also lists the n (which is of course zero since there are no calibration results yet), the counts of the four bins (which are also zero) and the values of A's three fractiles on the variable. It can be seen that the true value fell in the zeroth bin. This is taken into account in the posterior on variable A-2. Now n=1, y_0=1 and the probability of the zeroth bin is higher and the remaining bins proportionally lower, uniformly distributed between the fractiles. These results compare to the results pertaining to the same variable in appendix 2A, in which the bin distributions were plotted.

In 3A-2 again the true value falls in the zeroth bin. In variable A-3 we find indeed that the probability of the zeroth bin has increased even more. The general effect is that probability mass is slowly being shifted to the left to compensate for an at this point seemingly appearing overestimation tendencies of source A. In the legend of variable A-3 we see that, thanks to I2, hits are simply added, resulting in a y_0 of 2. This implies that all previous calibration indicators count just as much in the determination of the posterior. The effect of a previous hit never disappears, although it eventually gets drowned in an overabundance of other calibration indicators. Subsequent outcomes show that overestimation is not really typical of source A. Already at, say variable A-12, there is a significant amount of true values in the third bin accompanied by a shift of probability mass to the right in an effort to compensate for the now apparent underestimating tendencies of source A. It is also clear that probability mass is spread out more; the system is trying to compensate for A's overconfidence as evidenced by the relatively high counts of true values in the extreme bins. As we go along the variables we see, as we also noted in the previous chapter, that this tendency persists.

The results for source B are shown in appendix 3B. As can be seen and as was already noted in chapter 5, they are in the same comparable to those in appendix 2A, except that B is quite a bit more overconfident and a much stronger underestimator than A. Note that the actual information a source provides in the posterior is a function of both the calibration data and the location of the fractiles of a source. Although the calibration on variables A-41 and A-42 is hardly different, A is much more informative (with respect to e.g. B) on variable 41 than 42. Since his fractiles were very close to begin with, he remained informative in spite of a considerable "flattening" of his

density.

Finally, in appendix 3AB the posteriors are shown for the two-source case. Several remarks can be made. Firstly, the densities tend to be much more "peaked" than either of the single-source cases. Thus, the combination of two sources is more informative than any single source, as one might hope. Secondly, since we have six instead of three fractiles to work with, the densities have more degrees of freedom, for instance, in this case the densities can be bimodal. Variable AB-12 is a good example in which the sources apparently disagreed strongly on the location of the main probability mass. In other cases it arises due to a combination of A's and B's fractiles, like in variables AB-14, -16, -42, -51 and others.

7. EVALUATION OF THE DECISION AID

Of course these compensating tendencies of the posterior densities on the decision variables are highly interesting in themselves, but they do not answer the ultimate question: is the posterior better than the prior, or, in less technical terms: is the corrected information with this aid any better than just taking the sources' word?

To ask whether one probability distributions is better than another *pur sec* does not make much sense. What does make sense is to ask, a posteriori, whether a one decision or control action the operator took was better than another. To pursue this line thought, we have to structure the decision problem a little. To this end, let us suppose that the action the operator has to undertake is the choice of a value of the variable, that is, for all fifty-three variables he has to provide an estimate of its value. To do to this, the operator not only needs a distribution over the variable representing his knowledge, but also a loss function over the possible consequences representing his goals. In an estimation problem the consequences consist of misestimating the true value more or less severely. What will be considered more or less severe will be dictated by the particulars of his task. Often it will be just as bad to overestimate as to underestimate in which case the loss function is symmetrical around the true value. In other cases this will not be true; William Tell's loss function was probably not symmetric.

Therefore, to show rigorously that the decision aid is useful to the operator, we must show that, no matter what the operator's goals or values are, he is better off, in the long, using the output of the aid rather than listening to the sources. A trifle more technically said, we must show that no matter which loss function the operator employs, he accrues less (or at most equal) losses in the long run using the posterior than he does using the prior, no matter what his loss function is.

Since there is an infinity of possible loss functions it is an impossible task to show this empirically. What we can and will do is to select a family of typical loss functions and show that the aid performs better under these. We have chosen for a family of so-called bilinear loss functions which are sketched in figures 7.5-i to 7.5-iv. A bilinear loss function yields zero loss if the estimate is equal to the true value (t.v. in figure 7.5) of the variable and is linearly increasing as the estimate goes away from the true value. However, the right slope does not have to be equal to the left slope. The loss function in 7.5-i has a left slope which is nine times smaller than the right slope. This loss function strongly punishes overestimation; if the estimates based on the distribution are higher than the true value, more loss is accrued than when it is the same distance lower than the true value. Loss function 7.5-ii is actually symmetric, punishing overestimation and underestimation equally. The remaining two loss functions punish underestimation, 7.5-iv in the same amount as 5-i punishes underestimation.

Let l be the left slope and r the right slope. It is well known (see e.g. Berger [1985]) that the optimal or Bayesian estimate when using a bilinear loss function is an $l/(l+r)$-fractile of the operator's distribution. Thus, the Bayesian estimate when using the loss function 7.5-i is a .1-fractile of the operator's distribution. The symmetric loss function implies that a .5-fractile or the median of the posterior is the Bayesian estimate and so on. Both the quotient of the slopes and the estimation fractile are given in the legend of each loss function. This fact points out another advantage of the use of the bilinear loss function to evaluate the aid. One arbitrary factor in the graphs of the densities in appendix 3 are the bounds at the extremes of the densities, which were chosen by us. A fractile is insensitive to the particular choice of these bounds. Had we chosen for instance a quadratic loss function, then the optimal estimate is the mean which is certainly not insensitive to the choice of these bounds.

In order to compare the prior with the posterior, we can now calculate the Bayesian estimate of

each decision variable based on the prior and on the posterior, this being the appropriate fractile of these distributions. Based on the true value and the estimate we can calculate the losses the operator accrues for both the prior and the posterior distribution. In the long run, the losses accrued while using the posterior should be less than those accrued wile using the prior for all four loss functions.

In figures 7.1-i through 7.1-iv we show the cumulative losses for the four different loss functions when the operator only uses source A. The prior losses are shown by a dashed curve and the posterior losses by a solid curve. At the zeroth variable, that is before any estimation has been done, the losses are of course zero. After the first variable, the losses for the prior are equal to those for the posterior since prior and posterior are equal on this variable. At the second variable they differ. In figure 7.1-i, the prior accumulates much higher losses than the posterior. Evidently, the shift of probability mass to the left was a fortunate thing, which figures since variable A-2 also had the true value falling in the zeroth bin. Then the loss accrued on estimating the third variable is added to the losses already accrued and so on until the fifty-third variable at which both the total losses are almost equal. If we had more variables we could see whether this latter phenomenon is just a statistical fluctuation and that the two curves ultimately diverge with the dashed line the upper branch, as they should. On the whole the posterior losses lie below the prior losses. Thus, with this particular loss function, the operator is at any point better of using the aid than listening to the source, although the difference is not spectacular. The difference is even less spectacular when using the symmetric loss function, as evidenced by figure 7.1-ii. Large gains occur, however, when using the loss function which punishes underestimation as evidenced by figure 7.1-iv.

The reason that the improvement is small in the first three cases is that the first three loss function are sensitive to biases that the sources do not display. If source A is not an overestimator, then the aid cannot and should not correct for overestimation. Consequently the difference between the corrected and uncorrected versions are not spectacular when using a loss function that concentrates on overestimation. If the source's main bias is overconfidence, then a symmetric loss function will not bear these corrections out since median does not change if the density is "flattened". Of course, since A is a also an underestimator, there are improvements going on. However, it will take much more than fifty-three variables to see the two curves diverging significantly since the rate of divergence is so much smaller. What is important is that the aid did not do worse on neither of the loss functions, so one is never worse off using the aid.

Since B is quite a bit worse calibrated than A, we expect the gains with respect to his claims to be relatively higher then in A's case. A comparison between figure 7.1 and 7.2 shows that this is indeed so. Since B is very overconfident and quite an underestimator, both figure 7.2-i and 7.2-iv show strong divergence, especially so in 7.2-iv, the case in which underestimation is punished. The reason that 7.2-i's prior losses look as if they proceed in steps is that the aid is trying to shrink far to the left to avoid any overestimation, since this is punished heavily with that particular loss function. By doing so it might, on the average, be more often further away from the true value, but it avoids the extremely costly risk of overestimating the variable. The prior does this a couple of times (the steps) and, in the long run, ends up losing a lot more.

Figures 7.3-i through 7.3-iv shows the losses for the two-adviser case. Again, we find overall improvement with the most significant occurring again when the loss function punishing underestimation is used -figure 7.3-iv. Since the prior in the double-source case is different from the priors in the single-source cases, not much can be said about the comparison of the relative improvements with respect to the prior. What can be meaningfully compared are the posterior losses for the three cases. If combining information with this aiding system is to be useful, then the posterior losses in the two-source case should be lower than any of the single-source cases. In figure 7.4 these losses are shown for the three cases by the indicated lines for the same four loss functions. On the first three, the combination is nearly always better than any single source. Figure 7.4-iv is

an exception. It seems that the discrepancy, however, is due to a number of "bad" variables in the beginning, and that the system is trying to catch up. If we had more variables we could show whether this conjecture is indeed right. This highlights a problem of the multi-source case. In the two-source case we have sixteen bins instead of four. This means that we have quite a bit less calibration data available per bin to update on. Thus, if we have a several accidentally misleading calibration indicators in the beginning, it will take a lot longer for these to "wash out" in the incoming calibration results. We are currently studying methods to overcome these problems so that the aiding system can also be used when the number of sources and fractiles is large and n is still relatively small.

8. SUMMARY AND CONCLUSIONS

A computer aiding system designed to evaluate and combine sources of information for an operator in a supervisory control situation was presented. The aid is based on six assumptions, categorized into three form assumptions and three inference assumption, namely

F1. There exist a real valued random variable on the value of which the operator wishes to base his decisions, which he in turn wishes to take in a Bayesian manner.

F2. The sources give their information in terms of fractiles of their distribution on the variable in F1.

F3. The sources can be completely evaluated by a sequence of calibration indicators. These consist of a registration of the location of the true value of an arbitrary number of F1-like variables with respect to the F2-like information obtained on those variables.

I1. If the operator has no evidence concerning the characteristics of the sources (i.e., if the sequence in F3 has no elements), the operator starts off believing the sources.

I2. The operator considers the order in which the calibration indicators appear to be irrelevant, i.e., he considers calibration indicators to be exchangeable random variables.

I3. The operator is as unsure as possible about assumption I1.

These six assumption, if accepted by the operator, lead to an unique posterior distribution that represents the operator's beliefs on hearing the information from the sources and in light their characteristics as information sources. The assumptions are designed to model the typical situation in which the sources are much more knowledgeable than the operator, and in which the operator does not know much about the characteristics of the sources so he feels comfortable in relying on the automated procedure of measuring their characteristics by a sequence of calibration indicators.

An experiment was conducted with two different information sources advising the operator on fifty-three different decision variables. On every variable the previous variables were used as calibration indicators -their true values were considered to be meanwhile known.

The posterior distributions to these variables -the output of the decision aid- were compared under different loss functions to a policy of simply believing the sources, i.e. the prior. The aid was found to perform equal or better on every loss function. The aid performs equally well under loss functions that do not bear out the biases of the source, or when a source has no biases. Otherwise improvements are found. It was also found that the posterior given both sources was generally superior to the posteriors under any one source. A limiting factor in this evaluation is the small number of decision variables, making it unclear whether there were ultimately to be saving in accumulated loss in the cases in which the improvements were relatively small.

9. BIBLIOGRAPHY

Berger, J. O., "Statistical Decision Theory and Bayesian Analysis." 2nd ed., Springer, New York, 1985.

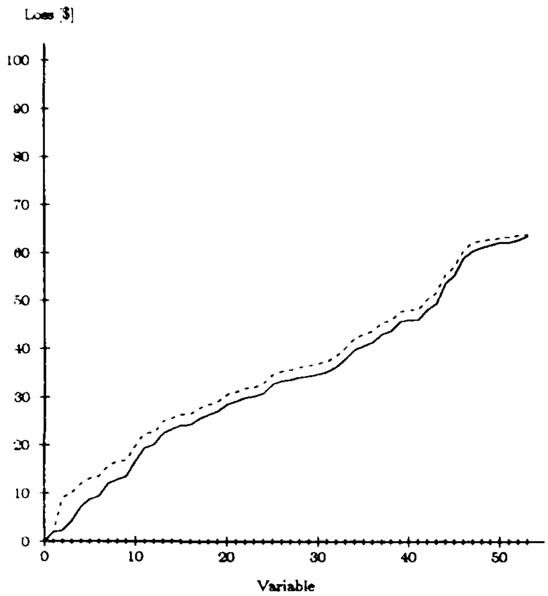
de Finetti, B., "Foresight: Its Logical Laws, Its Subjective Sources." Studies in Subjective Probability (Kyburg and Smokler eds.), Wiley, New York, 1964. (First appeared in 1937 in Ann. de l'Inst. H. Poincare in French).

de Finetti, B., "Theory of Probability." Vols 1 and 2, Wiley, New York, 1974, 1975.

Lichtenstein, S., Fischoff, B., Phillips, D., "Calibration of Probabilities: The State of the Art to 1980." In Kahneman et al., (eds) "Judgment under Uncertainty: Heuristics and Biases." Cambridge, Cambridge University Press, 1982.

Pfanzagl, J., "Theory of Measurement." Wiley, New York, 1986.

Savage, L., "The Foundations of Statistics." Wiley, New York, 1954.

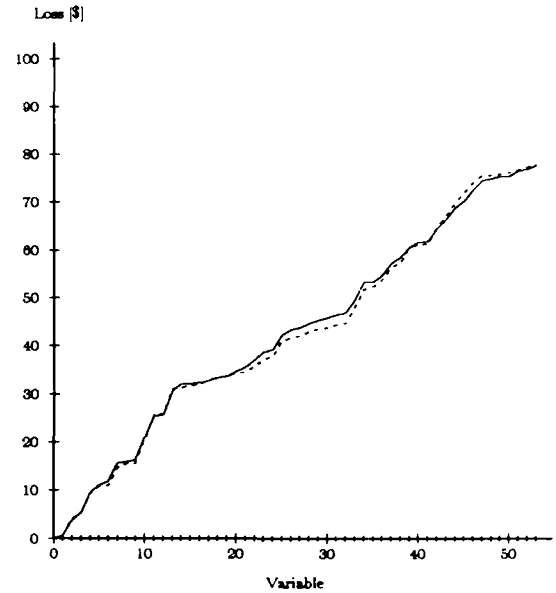


Source = A

Left slope : Right slope = 1 : 9

Estimation fractile = 0.1

Figure 7.1-i

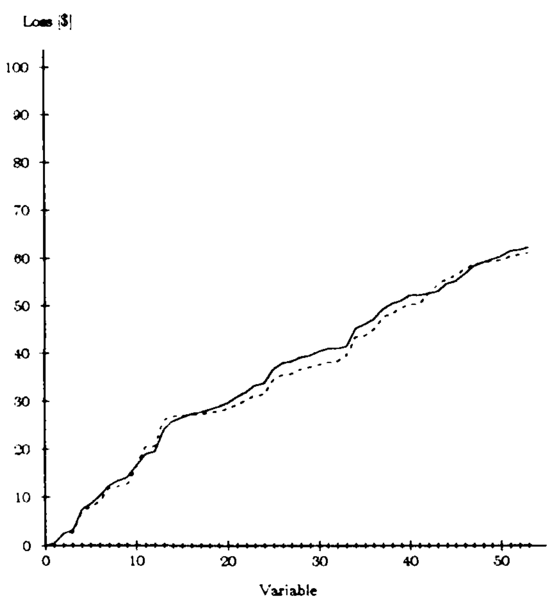


Source = A

Left slope : Right slope = 1 : 1

Estimation fractile = 0.5

Figure 7.1-ii

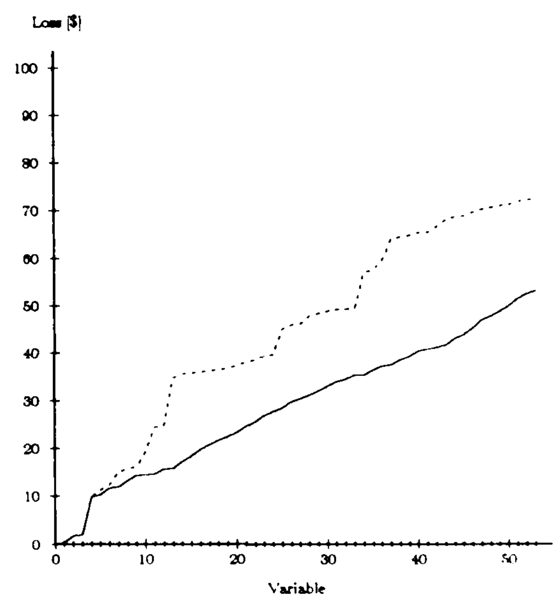


Source = A

Left slope : Right slope = 1 : 0.428571

Estimation fractile = 0.7

Figure 7.1-iii



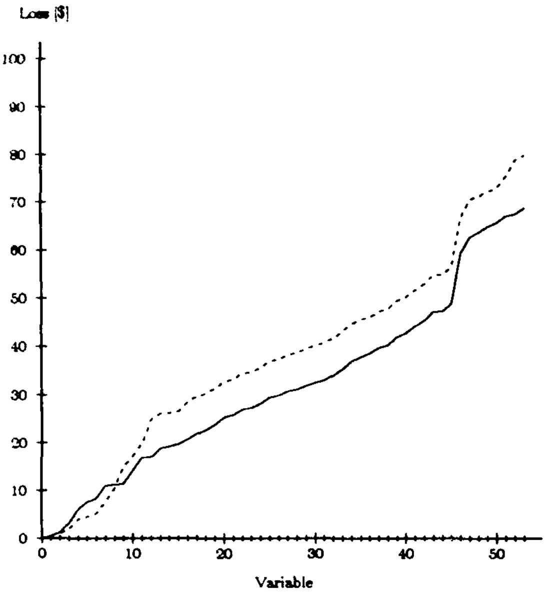
Source = A

Left slope : Right slope = 1 : 0.111111

Estimation fractile = 0.9

Figure 7.1-iv

Figure 7.1: Cumulative losses of the advice (dashed) and corrected advice (solid) of source A.

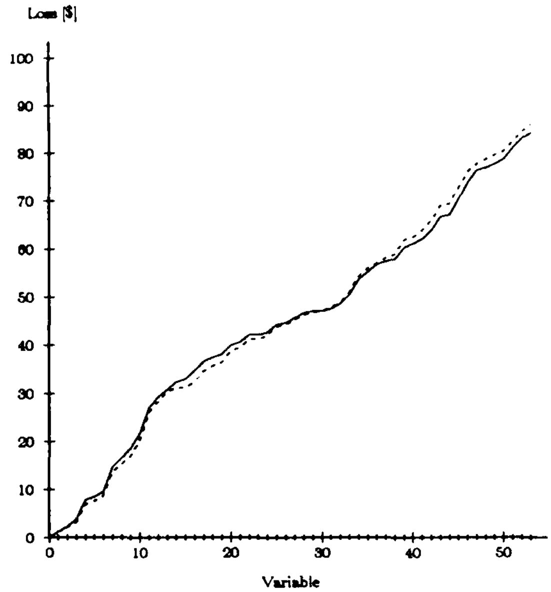


Source = B

Left slope : Right slope = 1 : 9

Estimation fractile = 0.1

Figure 7.2-i

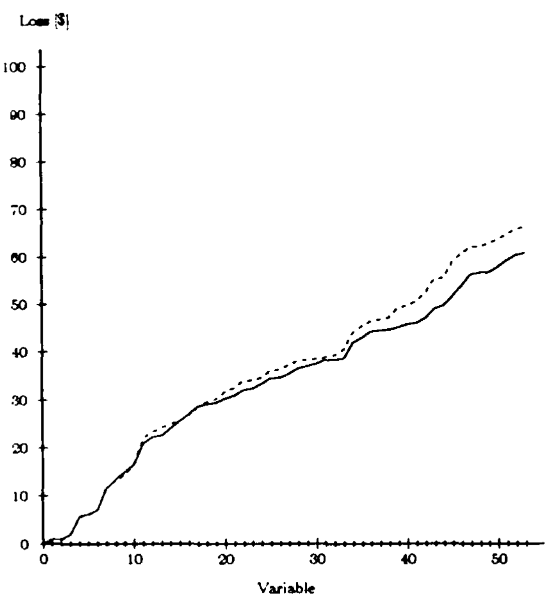


Source = B

Left slope : Right slope = 1 : 1

Estimation fractile = 0.5

Figure 7.2-ii

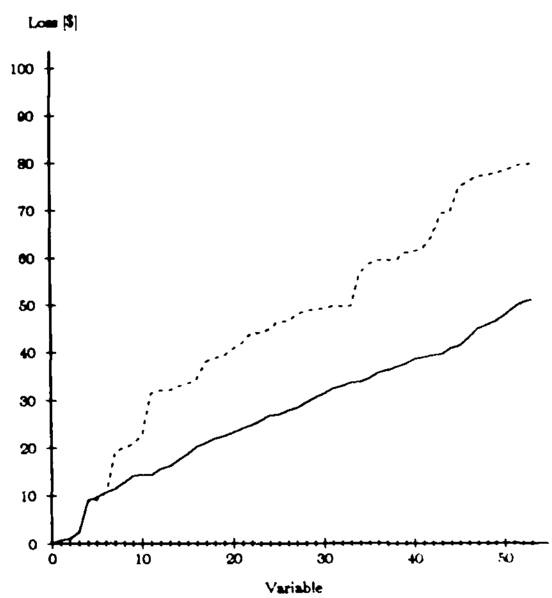


Source = B

Left slope : Right slope = 1 : 0.428571

Estimation fractile = 0.7

Figure 7.2-iii



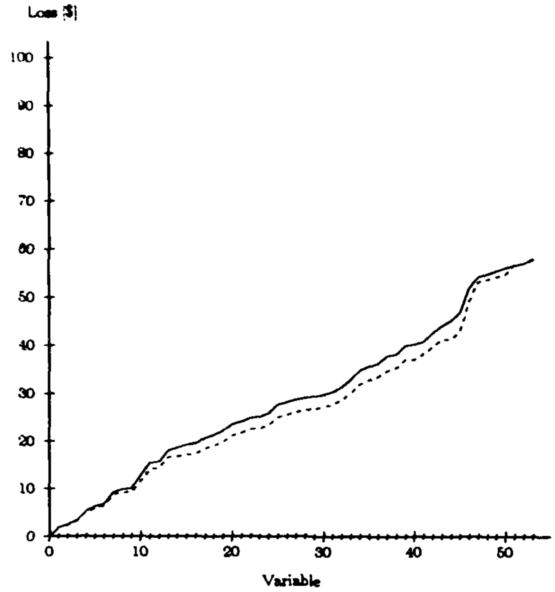
Source = B

Left slope : Right slope = 1 : 0.111111

Estimation fractile = 0.9

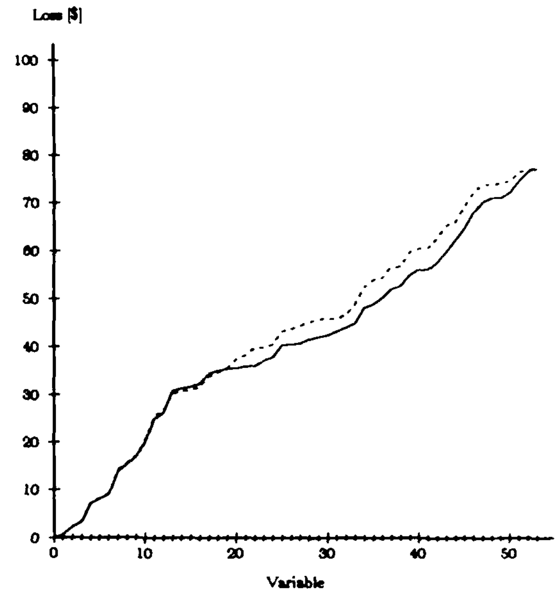
Figure 7.2-iv

Figure 7.2: Cumulative losses of the advice (dashed) and corrected advice (solid) of source B.



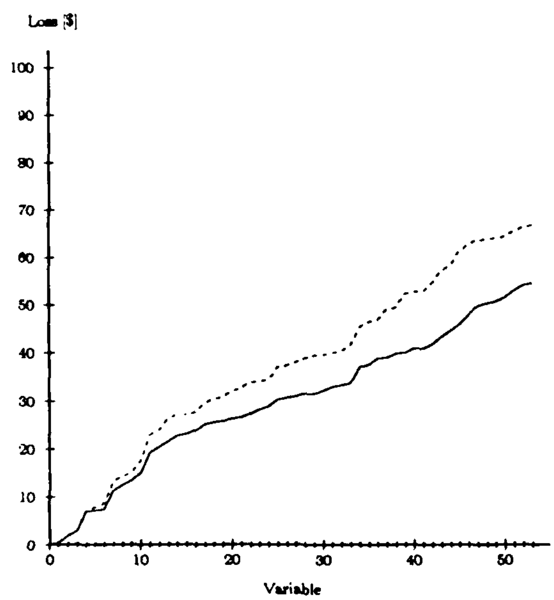
Source = AB
Left slope : Right slope = 1 : 9
Estimation fractile = 0.1

Figure 7.3-i



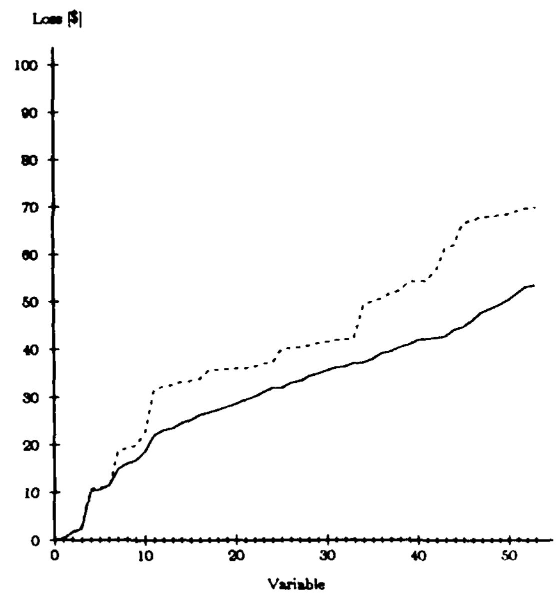
Source = AB
Left slope : Right slope = 1 : 1
Estimation fractile = 0.5

Figure 7.3-ii



Source = AB
Left slope : Right slope = 1 : 0.428571
Estimation fractile = 0.7

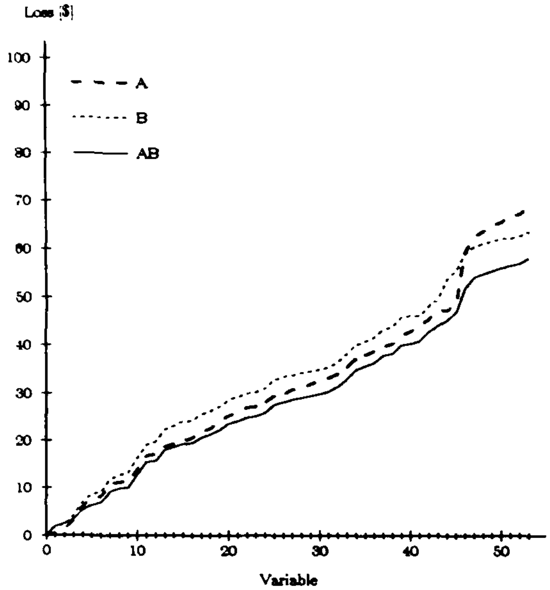
Figure 7.3-iii



Source = AB
Left slope : Right slope = 1 : 0.111111
Estimation fractile = 0.9

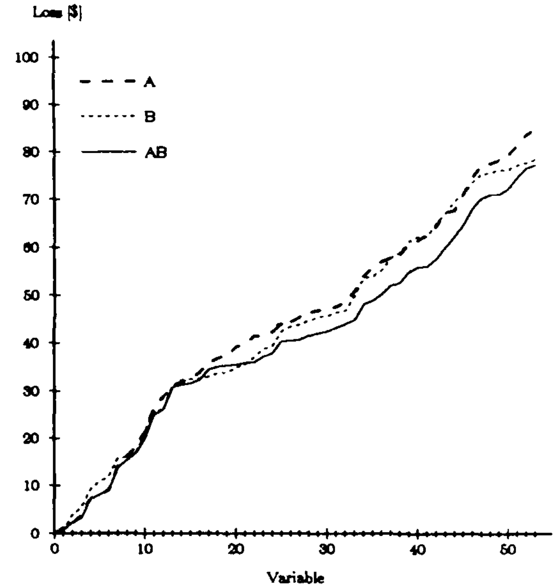
Figure 7.3-iv

Figure 7.3: Cumulative losses of the advice (dashed) and corrected advice (solid) of sources A and B together.



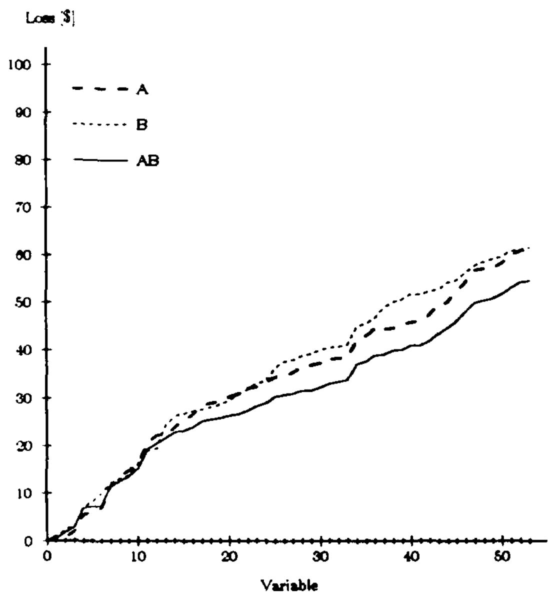
Left slope : Right slope = 1 : 9
Estimation fractile = 0.1

Figure 7.4-i



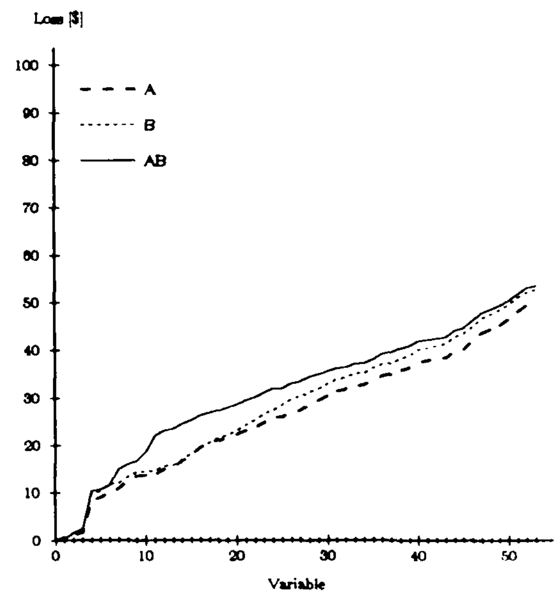
Left slope : Right slope = 1 : 1
Estimation fractile = 0.5

Figure 7.4-ii



Left slope : Right slope = 1 : 0.428571
Estimation fractile = 0.7

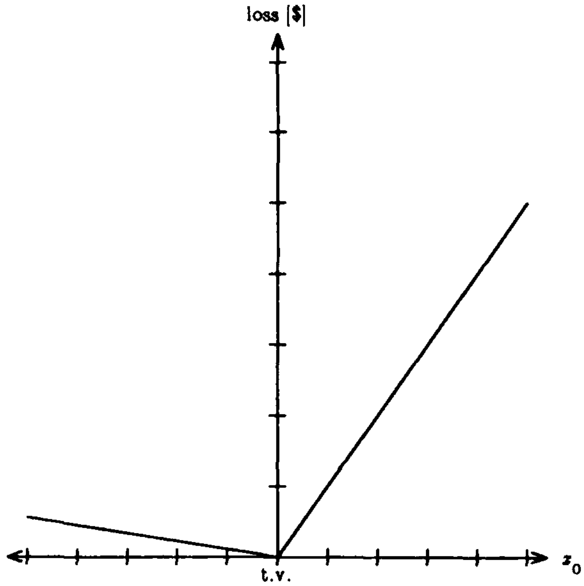
Figure 7.4-iii



Left slope : Right slope = 1 : 0.111111
Estimation fractile = 0.9

Figure 7.4-iv

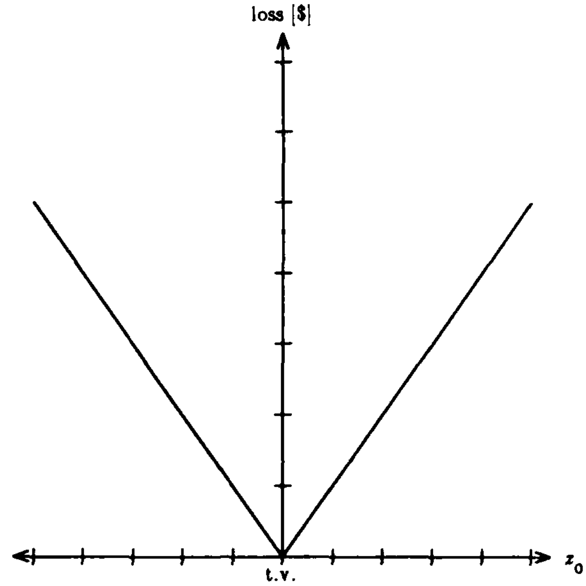
Figure 7.4: Cumulative losses of the corrected advice of sources A, B and A and B combined.



Left slope : Right slope = 1 : 0.111111

Estimation fractile = 0.1

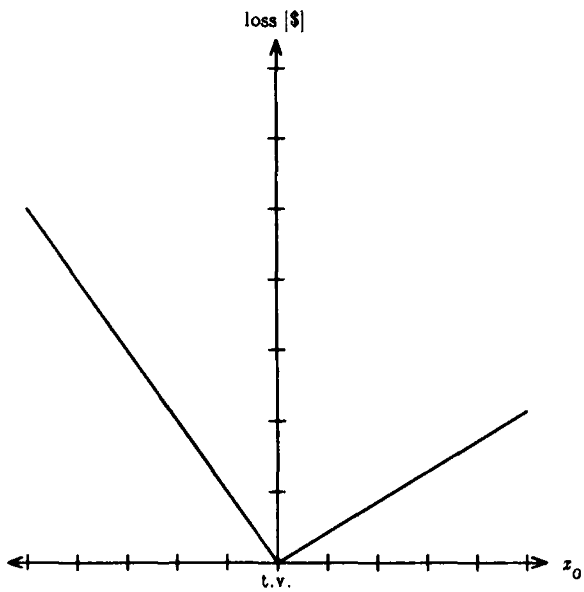
Figure 7.5-i



Left slope : Right slope = 1 : 1

Estimation fractile = 0.5

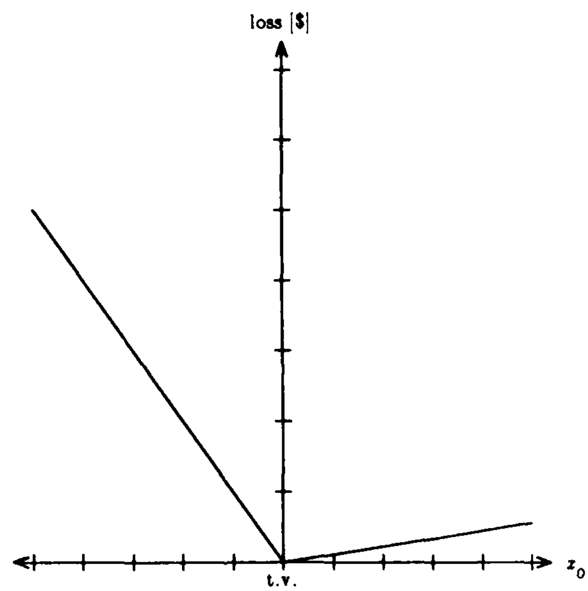
Figure 7.5-ii



Left slope : Right slope = 1 : 2.33333

Estimation fractile = 0.7

Figure 7.5-iii



Left slope : Right slope = 1 : 9

Estimation fractile = 0.9

Figure 7.5-iv

Figure 7.5: Bilinear loss functions with various slope ratios.

9. BIBLIOGRAPHY

Berger, J. O., "Statistical Decision Theory and Bayesian Analysis." 2nd ed., Springer, New York, 1985.

de Finetti, B., "Foresight: Its Logical Laws, Its Subjective Sources." Studies in Subjective Probability (Kyburg and Smokler eds.), Wiley, New York, 1964. (First appeared in 1937 in Ann. de l'Inst. H. Poincare in French).

de Finetti, B., "Theory of Probability." Vols 1 and 2, Wiley, New York, 1974, 1975.

Lichtenstein, S., Fischoff, B., Phillips, D., "Calibration of Probabilities: The State of the Art to 1980." In Kahneman et al., (eds) "Judgment under Uncertainty: Heuristics and Biases." Cambridge, Cambridge University Press, 1982.

Pfanzagl, J., "Theory of Measurement." Wiley, New York, 1986.

Savage, L., "The Foundations of Statistics." Wiley, New York, 1954.

11. Appendix 1

(1) The Lamborghini Countach LP 500 S holds the world record of the highest *tested* speed. It was kept out of the "red zone", i.e. at below 7000 rpm. What was this speed?

(2) What is the highest speed attained by a rocket-powered ice sled?

(3) How much power does the Grand Coulee, the most powerful power station in the world, currently produce?

(4) Ford's 15-cc UFO2 (Ultimate Fuel Optimiser) holds the record for the lowest gasoline consumption. It used the burn/coast method and drove between 13 and 22 mph. What was this record fuel consumption?

(5) What was the longest recorded distance for driving on 2 wheels in a regular production line car?

(6) The longest recorded skidmarks on a public road were left by a Jaguar involved in an accident on the M1 (English freeway) near Luton, England. Evidence given in court indicate a speed "in excess of 100 mph" before the application of the brakes. How long were these skidmarks?

(7) The fastest circumnavigation of the world was done by two Canadians in a Volvo 245. How long did it take them to do this?

(8) Two people from Photovoltaic Power Systems in Calif. attained the highest speed ever for solar-powered vehicles with their "Sunrunner". What was their speed?

(9) Evidently there exist a practical upper limit for the size of tires. Goodyear have come the closest to that by manufacturing tires for giant dump trucks. What is their diameter?

(10) how much do they weigh

(11) and how much do they cost?

(12) What is the world speed record for human powered (i.e. one human) vehicles?

(13) What is the tallest unicycle ever mastered? It was ridden a distance of 376 ft?

(14) What is the highest speed attained by a railed vehicle?

(15) The longest railroad line is the Trans-Siberian, running between Moscow and Nakhodka. How long is it?

(16) The highest recorded takeoff weight of any aircraft was the case of a Boeing 747-200B Jumbo jet during certification tests of its Pratt & Whitney JT9D-7Q engines. How much was that?

(17) The French hold the altitude record for helicopters set by a 315 B Lama, by Aerospatiale SA. What is that height?

(18) Winzen Research Inc. in Minnesota built the largest balloon ever. What was its volume?

(19) The first major tidal power station is the "Usine Maremotrice de la Rance", opened in 1966 at the Rance estuary in the gulf of St. Malo, Brittany, France. It has a net annual output of 544 million kWhours. How long is its barrage?

(20) and how many turbo-alternaters does it contain?

(22) California boasts the world's largest solar power plant. How much does it produce?

(21) Babcock & Wilcox Co. designed the largest boilers ever. How many lbs of steam per hours are evaporated in one such a boiler?

(23) The largest catalytic cracker is the Exxon's Baywater Refinery plant at Linden, NJ. What is its fresh feed rate in gallons per day?

(24) The largest nuts ever made, the so-called "Pilgrim Nuts", manufactured in England, were used on the columns of a large forging press. What was their diameter?

(25) and what was their thread?

(26) The most powerful cranes in the world are two cranes operated by the dutch company 'Heerema'. They are used in tandem to set platforms in place on the continental shelf off the coast of Holland. How much can these cranes lift (in tandem)?

(27) The fastest printer is the Radiation Inc electro-sensitive system at the Lawrence Livermore Lab, Calif. The speed is attained by controlling electronic pulses through chemically impregnated recording paper which is rapidly moving under closed fixed styli. How many lines (120 alphanumeric characters) can it type per minute?

(28) The most accurate time keeping devices are the twin atomic hydrogen masers installed at the US Naval Research Lab. They are based on the frequency of the hydrogen atom's transition period. What accuracy does this enable them to have?

(29) The Olsen clockwork, installed in the Copenhagen town hall is the most accurate mechanical clock. How accurate is this clockwork?

(30) The world's most accurate computer to date is the CRAY-1, designed by Seymour R. Cray of Cray research inc., Minneapolis. What is its clockperiod?

(31) how many bytes of main memory does it have?

(32) and how many floating point operations per second does it attain?

(33) The thinnest wristwatch ever made is, of course, Swiss. How thick is it?

(34) The highest known prime number was discovered on the previously mentioned Cray-1. How high is it?

(35) Yasumasu Kanada of Japan has calculated π to the highest number of decimal places. Part of his expansion has been published in what is known as the world's most boring 800 pages. How long was his original expansion? (By the way, the most *inaccurate* version was made by the General Assembly of Indiana, who enacted in House Bill #246 that π was *de jure* 4.)

(36) The most accurate balance is the Sartorius Model 4108, manufactured in Goettingen (D). To which accuracy can it weigh objects of up to .5 grams?

(37) The $13 million Large Optics Diamond Turning Machine at Lawrence Livermore labs can make the finest cuts in the world. Estimate how many times it was able to sever a human hair lengthwise.

(38) The hottest flame that can be produced is from carbon subnitrite (C^4N^2). Which temperature is it calculated to reach at standard conditions?

(39) The Grand Coulee also possesses the largest hydraulic turbines, installed by Allis-Chalmers. At how many megawatts are these turbines rated?

(40) The lowest coefficient of static and dynamic (being the same in this case) friction of any solid is the case of polytetrafluorethylene ($[C_2F_4]_n$) or PTFE. It is manufactured by du Pont and marketed as Teflon. Estimate this coefficient.

(41) Both the strongest and the weakest magnetic fields were achieved at the Francis Bitter labs here at MIT. The weak field is used for research concerning the very weak magnetic fields generated in the heart and brain. What is the strongest field

(42) and the weakest?

(43) At MIT we also hold the record for highest note yet attained. Such tones are produced by a laser beam striking a sapphire crystal. What is its frequency?

(44) The highest man-made temperatures occur in the center of a thermonuclear fusion bomb, and are of the order of 300-400 million K. The highest *controllable* temperature have been achieved at Princeton's Plasma Physics Lab. in the fusion research Princeton Large Torus. The lowest temperature was reached in a two-stage nuclear demagnetization cryostat at Espoo in Finland. What was this largest?

(45) and lowest

(46) The Naval Research Lab. in Washington DC, holds the record for the highest velocity any solid visible object has ever attained, by projecting a plastic disc. What was this velocity?

(47) The Eiffel tower, designed by Alexandre Eiffel for the Paris exhibition, has a length of 985 ft 11 in tall. Estimate its maximum sway in high winds.

(48) The longest bridge span is the main span of the Humber Estuary bridge in England, at 4,626 ft. The bridge's two towers are 533ft 15/8in tall from datum and are brought out of parallel to allow for the curvature of the earth. Estimate how many inches they are brought out of parallel.

(49) The greatest of the roman aquaducts was the Aquaduct of Carthage (now in Tunesia), which ran 87.6 miles. It was built during the reign of Hadrianus (117-138 AD). How much water per day did it originally supply to the city of Carthage.

(50) In 1985, at MIT, Thomas Stockebrand intended to prove that a railway system could be built to operate at velocities greater than the speed of sound. Using a working model of a supersonic subway tunnel, before world leaders in the fields, pneumatically pulled a table tennis ball through 950 ft of small diameter pipe. At which speed did the ball go?

(51) The tallest unsupported flagpole stands at Chula Vista, Calif, and flies the Stars and Stripes. What is its length (above the ground)?

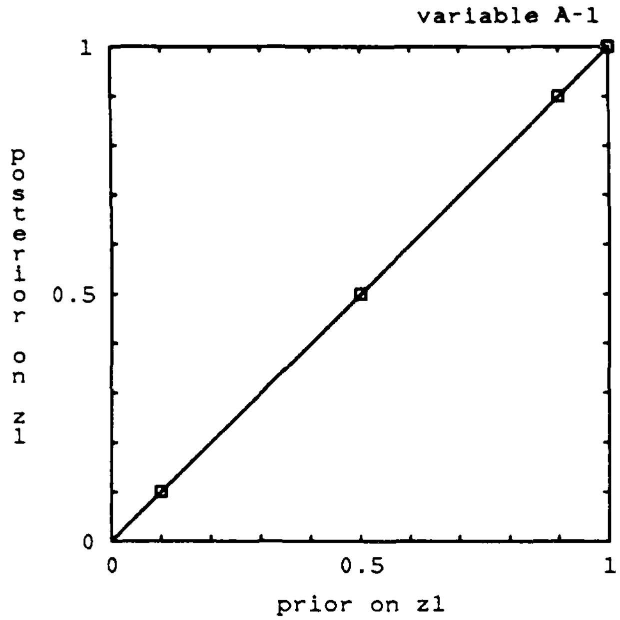
(52) The fastest speed at which any human has traveled was attained by the crew of the Apollo X when the Command Module reached its maximum speed on the trans earth return flight at an altitude of 400,000 ft. What was this speed?

(53) The record ocean descent was achieved in the Challenger Deep of the Marianes Trench, 250 miles southwest of Guam, when the Swiss-built US Navy Bathyscaphe "Trieste" reached the ocean bed. At what depth was this?

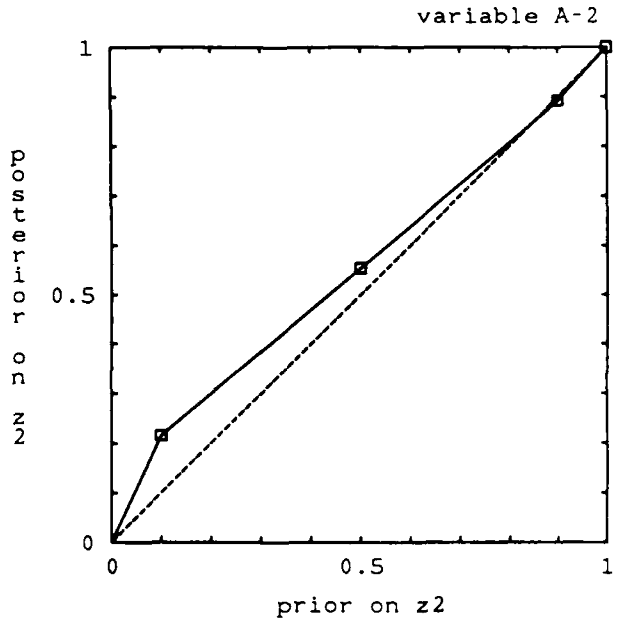
11. Appendix 2

11.1 Appendix 2A

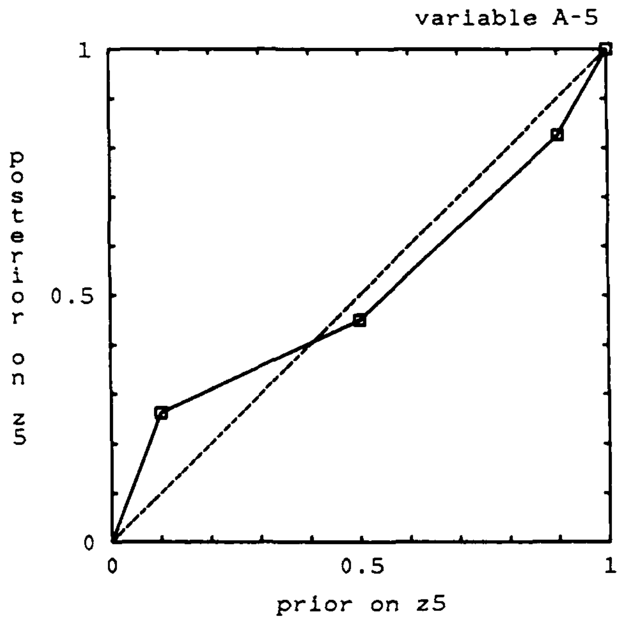
Calibration plots of source A



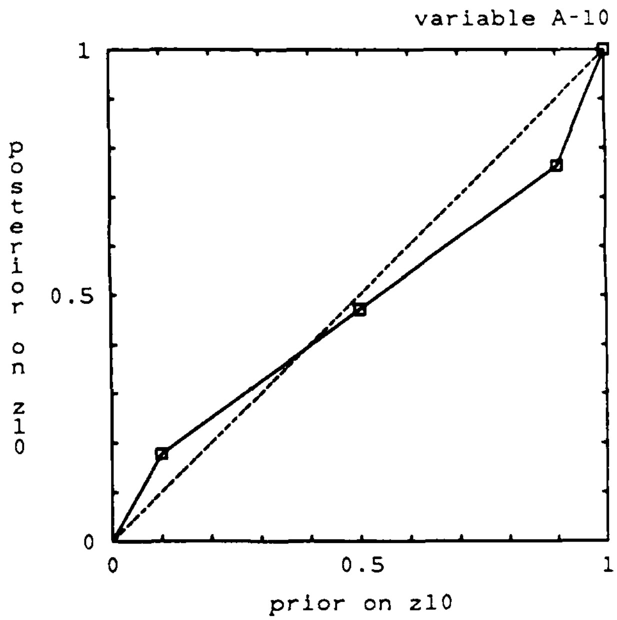
y0 = 0 y1 = 0 y2 = 0 y3 = 0



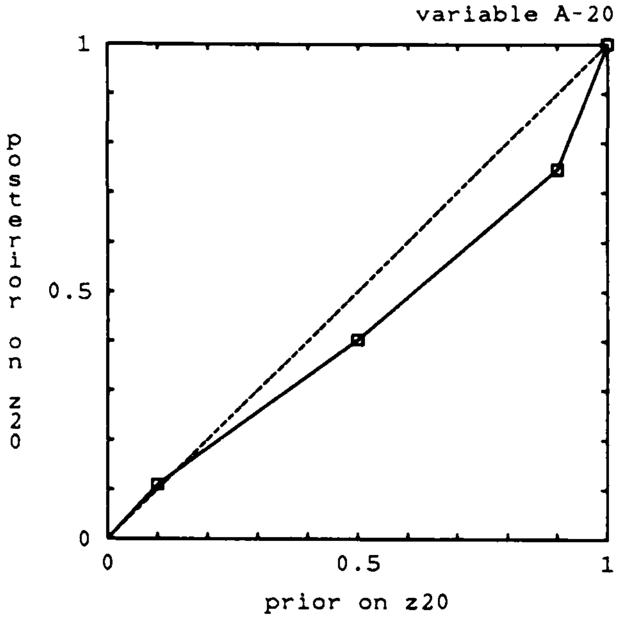
y0 = 1 y1 = 0 y2 = 0 y3 = 0



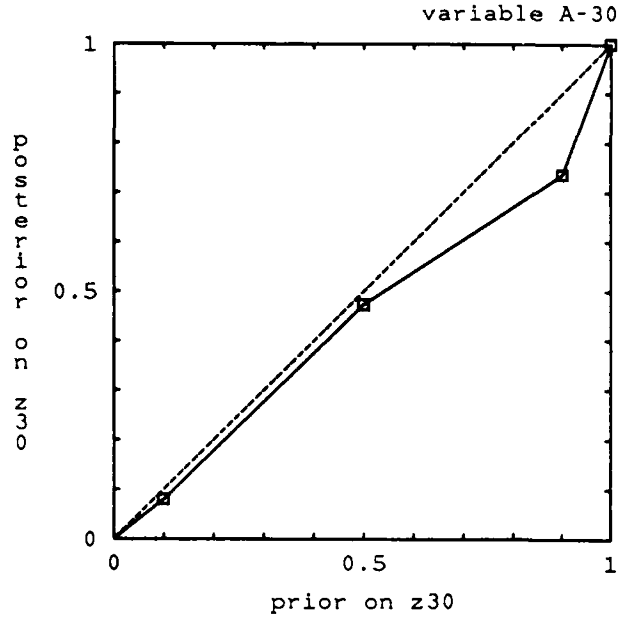
y0 = 2 y1 = 0 y2 = 1 y3 = 1



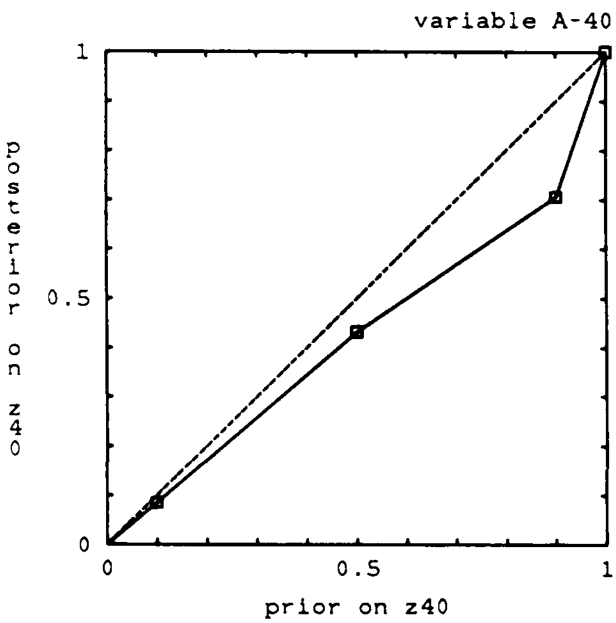
y0 = 2 y1 = 2 y2 = 2 y3 = 3



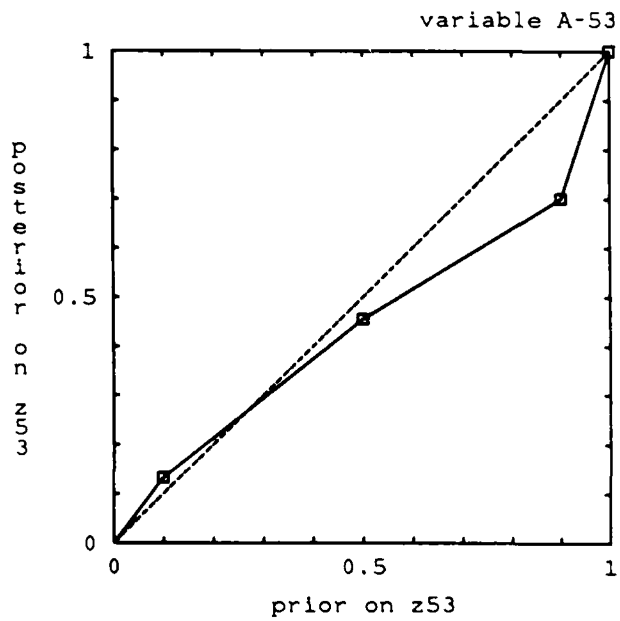
y0 = 2   y1 = 5   y2 = 6   y3 = 6



y0 = 2   y1 = 11   y2 = 7   y3 = 9



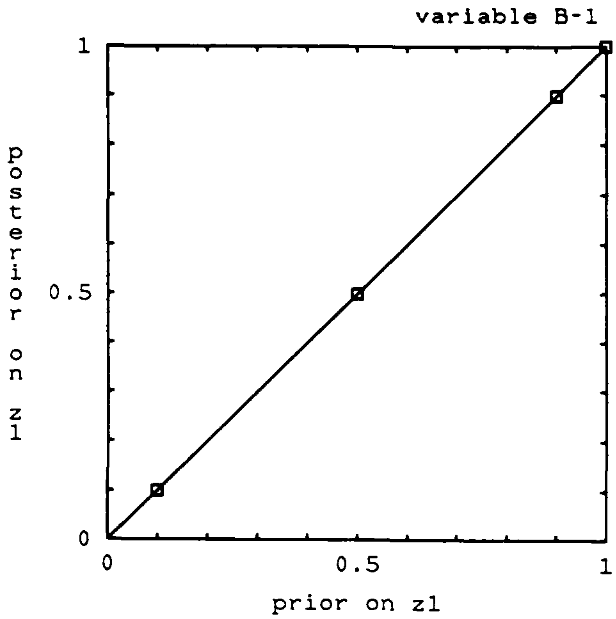
y0 = 3   y1 = 13   y2 = 10   y3 = 13



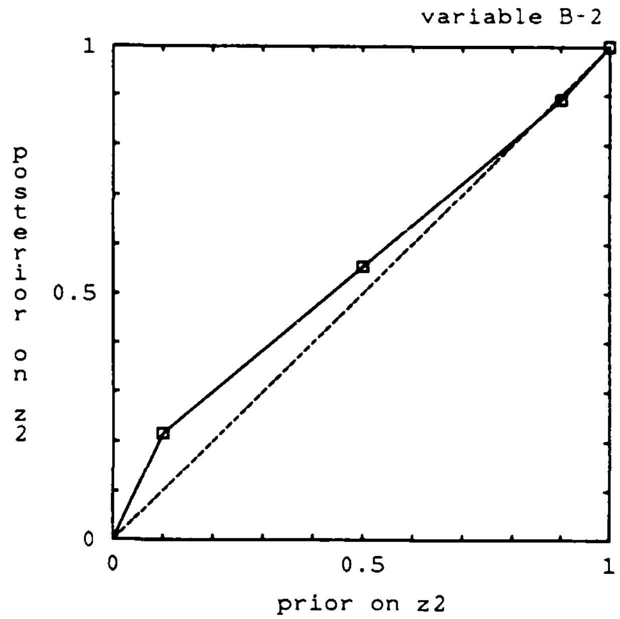
y0 = 7   y1 = 16   y2 = 12   y3 = 17

11.2 Appendix 2B

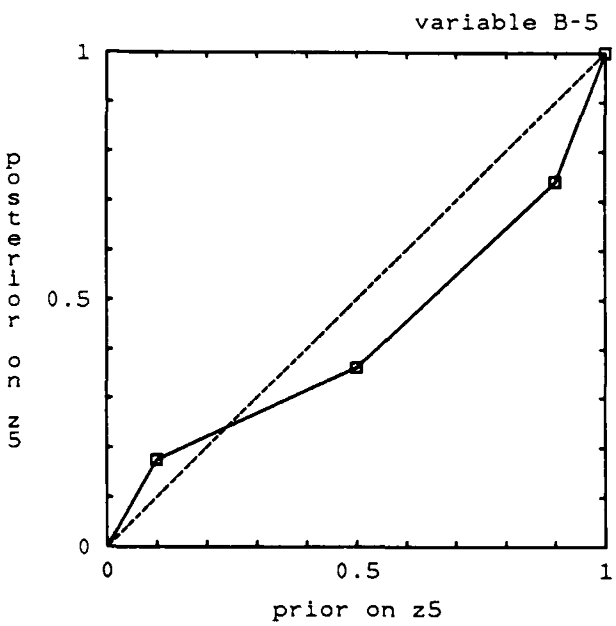
Calibration plots of source B



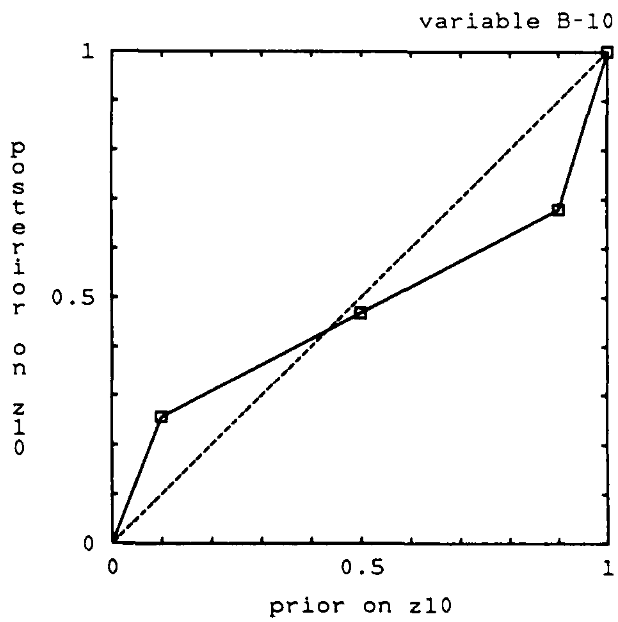
y0 = 0 y1 = 0 y2 = 0 y3 = 0



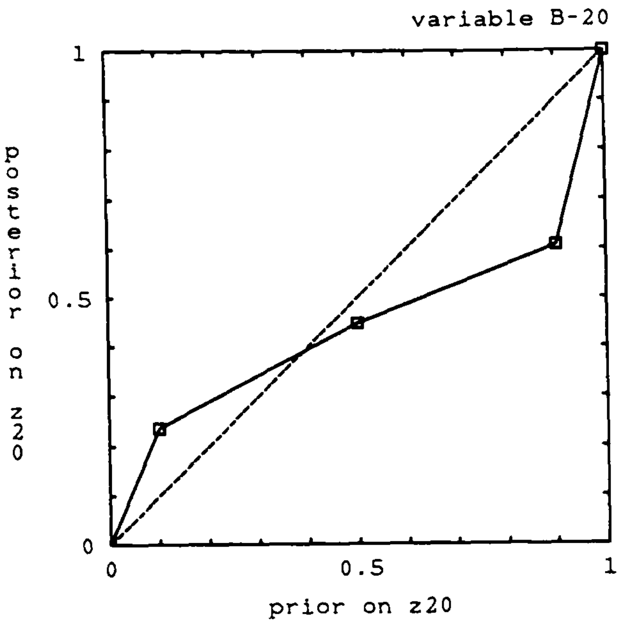
y0 = 1 y1 = 0 y2 = 0 y3 = 0



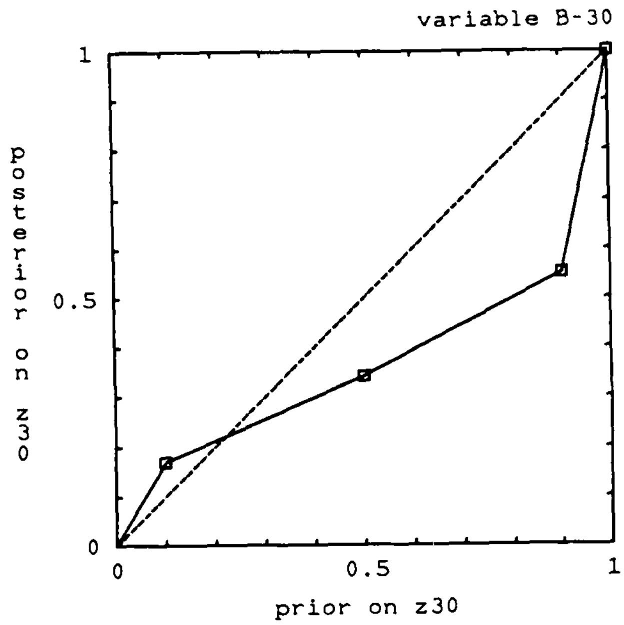
y0 = 1 y1 = 0 y2 = 1 y3 = 2



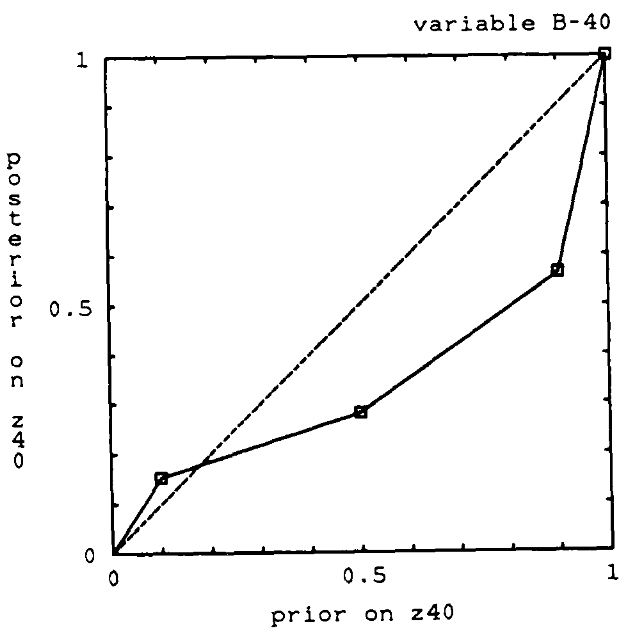
y0 = 3 y1 = 1 y2 = 1 y3 = 4



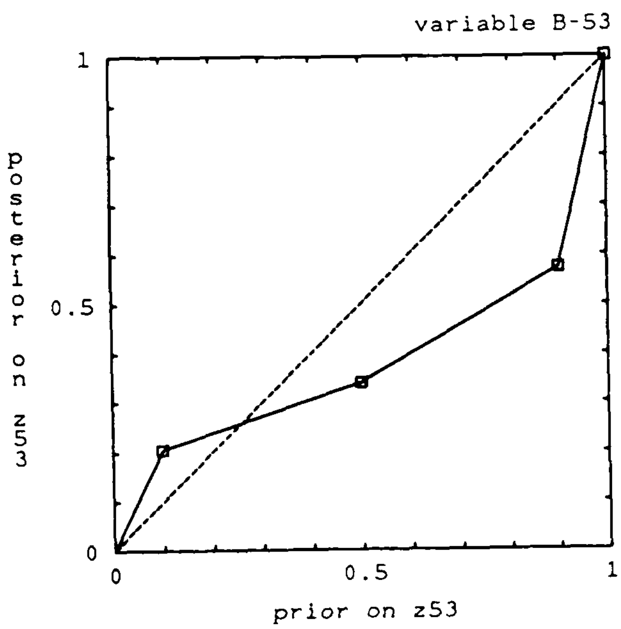
y0 = 5   y1 = 3   y2 = 2   y3 = 9



y0 = 5   y1 = 4   y2 = 5   y3 = 15



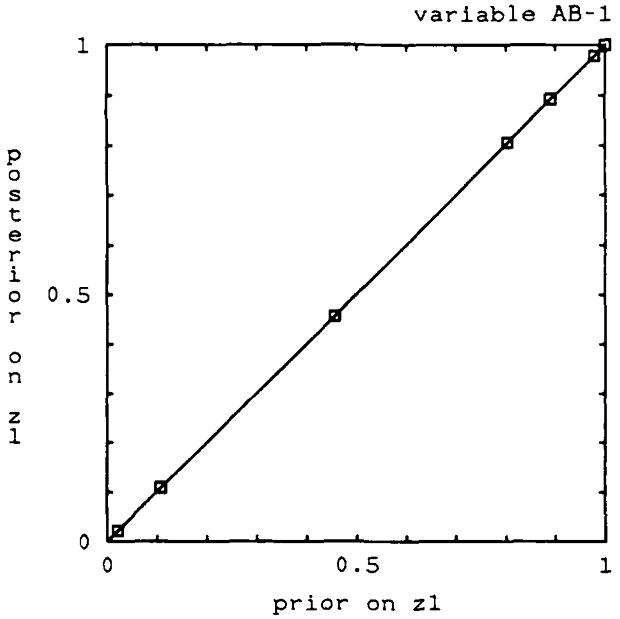
y0 = 6   y1 = 4   y2 = 10   y3 = 19



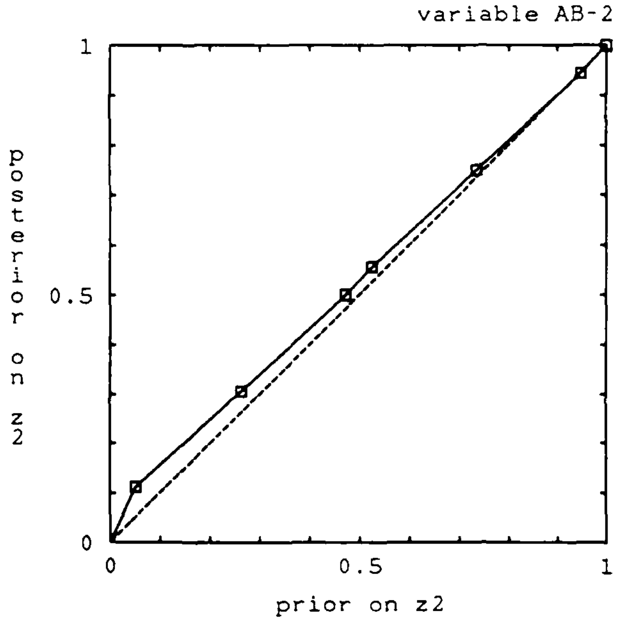
y0 = 11   y1 = 6   y2 = 11   y3 = 24

11.3 Appendix 2AB

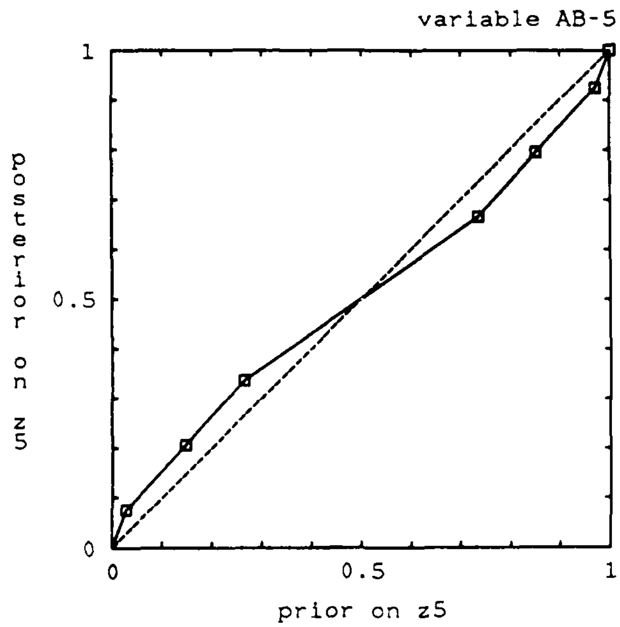
Calibration plots of sources A and B together



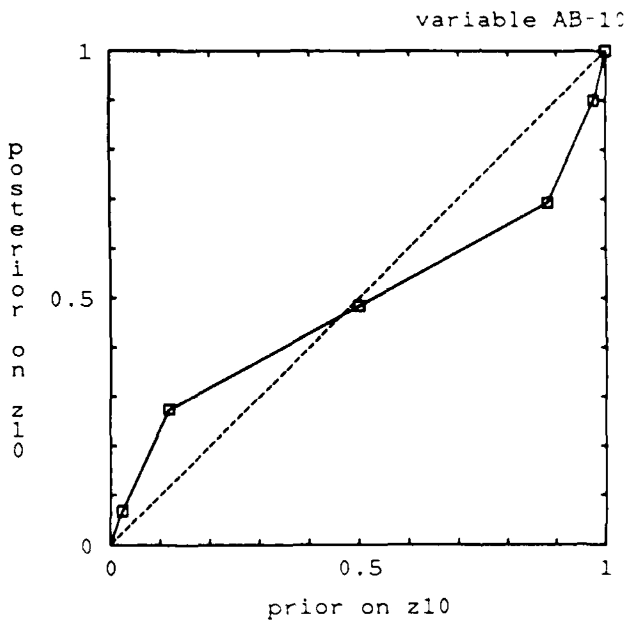
y00 = 0   y01 = 0   y11 = 0   y21 = 0
y31 = 0   y32 = 0   y33 = 0



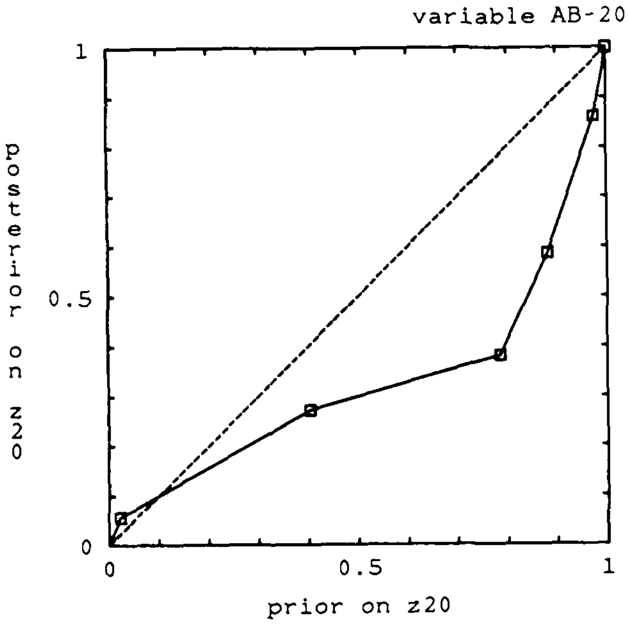
y00 = 1   y01 = 0   y02 = 0   y03 = 0
y13 = 0   y23 = 0   y33 = 0



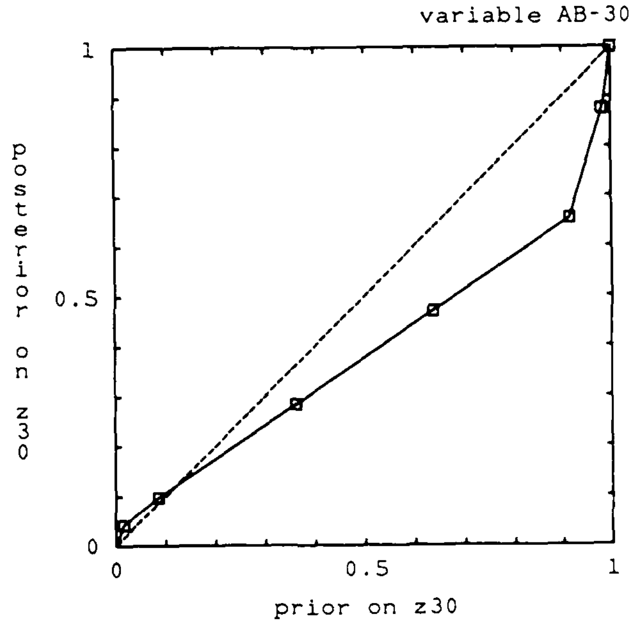
y00 = 1   y10 = 0   y20 = 0   y21 = 0
y31 = 0   y32 = 0   y33 = 1



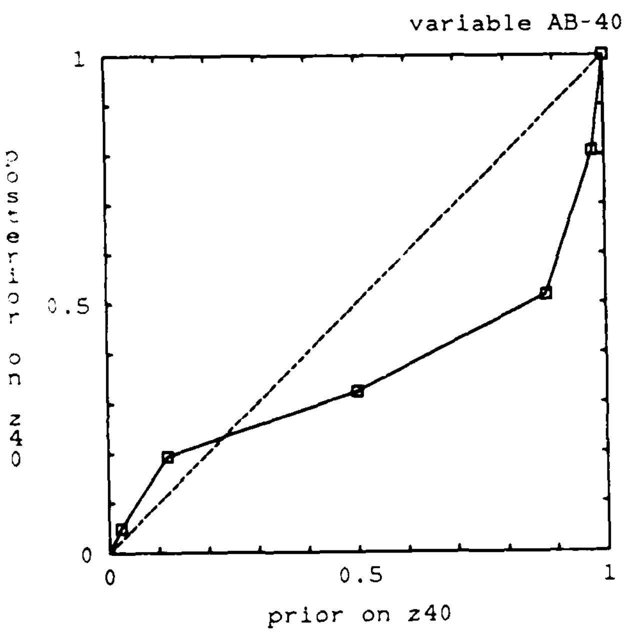
y00 = 1   y10 = 1   y11 = 0   y21 = 0
y22 = 0   y32 = 1   y33 = 2



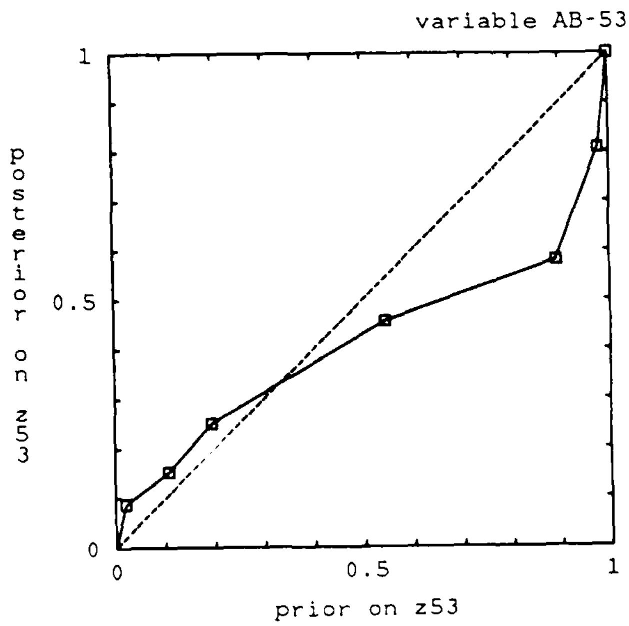
y00 = 1   y01 = 0   y11 = 1   y12 = 0
y13 = 2   y23 = 3   y33 = 4



y00 = 1   y01 = 0   y11 = 1   y12 = 1
y22 = 1   y23 = 3   y33 = 5



y00 = 1   y10 = 2   y11 = 1   y12 = 3
y22 = 2   y23 = 5   y33 = 7

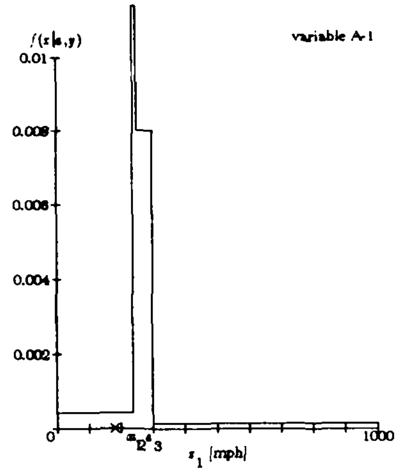


y00 = 4   y01 = 1   y02 = 2   y12 = 4
y22 = 2   y23 = 6   y33 = 10

11. Appendix 3

11.1 Appendix 3A

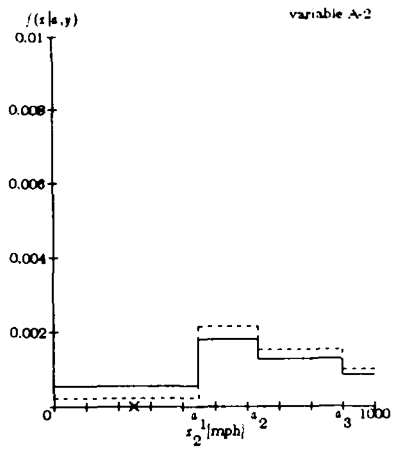
Advice (dashed) and corrected advice (solid) of source A



$n = 0$ true value = 179.9

$y_0 = 0$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

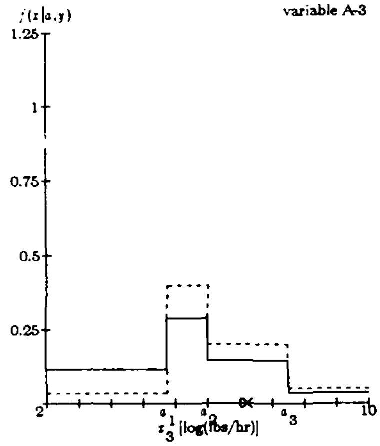
$a_1 = 238$ $a_2 = 250$ $a_3 = 300$



$n = 1$ true value = 247.93

$y_0 = 1$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

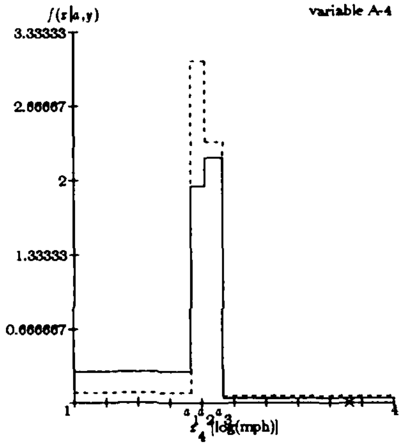
$a_1 = 450$ $a_2 = 636$ $a_3 = 900$



$n = 2$ true value = 6.97

$y_0 = 2$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

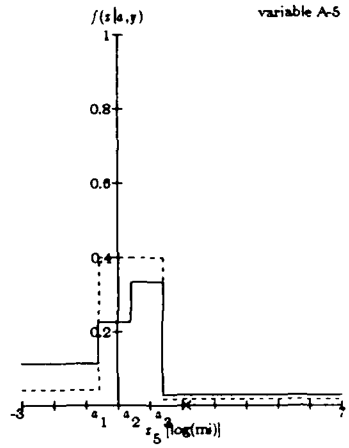
$a_1 = 5$ $a_2 = 6$ $a_3 = 8$



$n = 3$ true value = 3.58

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 0$

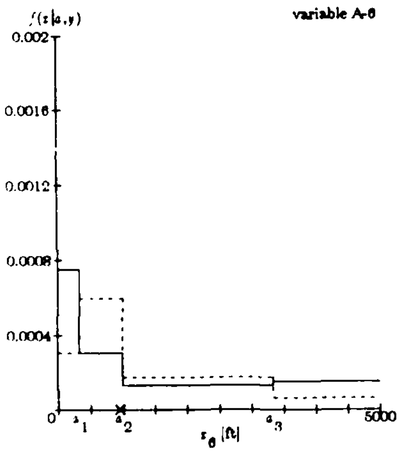
$a_1 = 2.1$ $a_2 = 2.23$ $a_3 = 2.4$



$n = 4$ true value = 2.114

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 1$

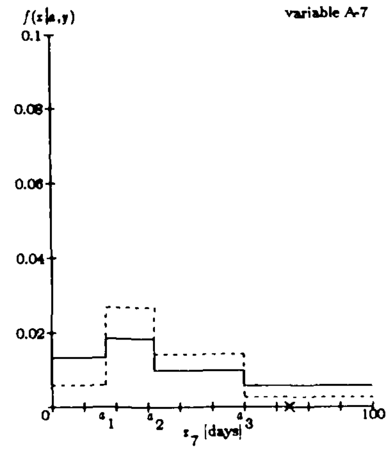
$a_1 = -0.6$ $a_2 = 0.4$ $a_3 = 1.4$



$n = 5$ true value = 950

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 2$

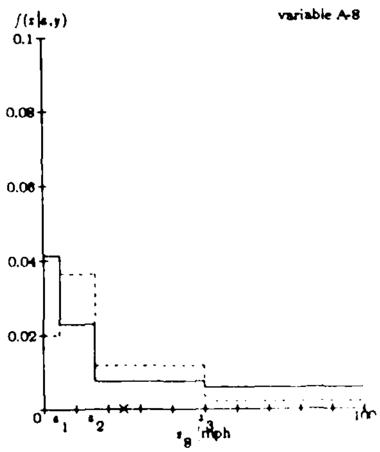
$a_1 = 330$ $a_2 = 1000$ $a_3 = 3340$



$n = 6$ true value = 74

$y_0 = 2$ $y_1 = 1$ $y_2 = 1$ $y_3 = 2$

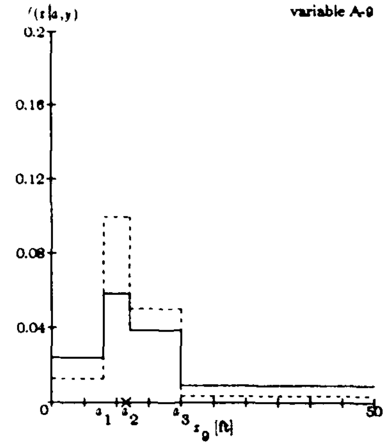
$a_1 = 17$ $a_2 = 32$ $a_3 = 60$



$n = 7$ true value = 24.74

$y_0 = 2$ $y_1 = 1$ $y_2 = 1$ $y_3 = 3$

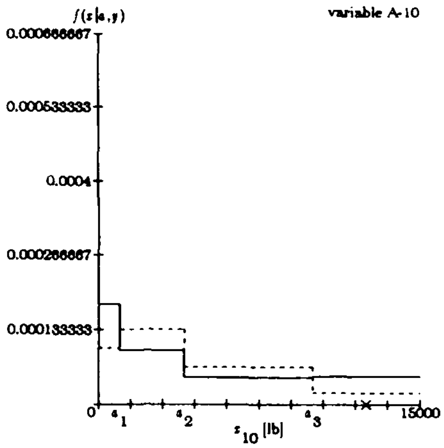
$a_1 = 5$ $a_2 = 16$ $a_3 = 50$



$n = 8$ true value $= 11.5$

$y_0 = 2$ $y_1 = 1$ $y_2 = 2$ $y_3 = 3$

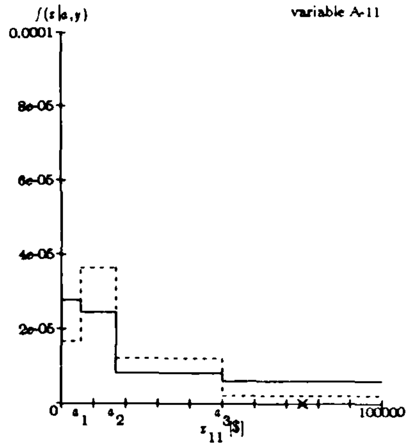
$a_1 = 8$ $a_2 = 12$ $a_3 = 20$



$n = 9$ true value $= 12500$

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 3$

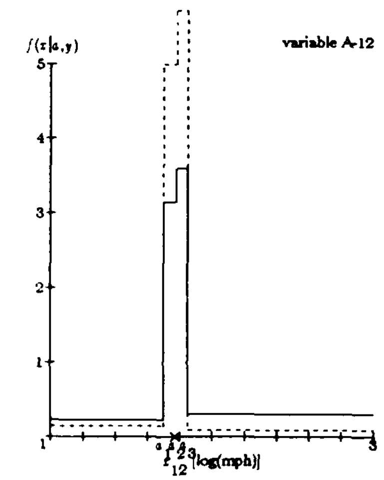
$a_1 = 1000$ $a_2 = 4000$ $a_3 = 10000$



$n = 10$ true value $= 75000$

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 4$

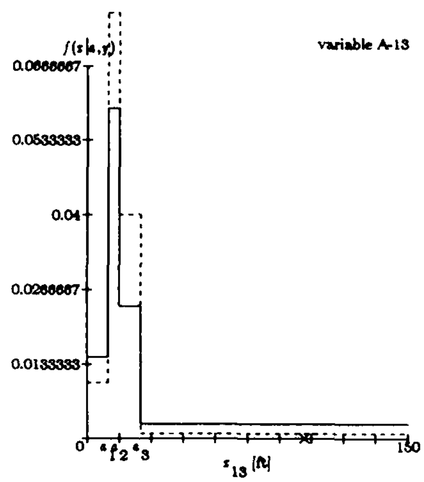
$a_1 = 6000$ $a_2 = 17000$ $a_3 = 50000$



$n = 11$ true value $= 1.77$

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 5$

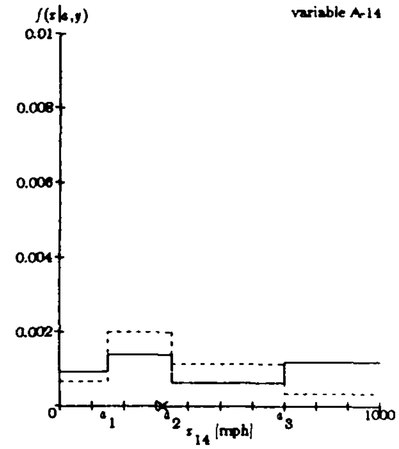
$a_1 = 1.7$ $a_2 = 1.78$ $a_3 = 1.85$



$n = 12$ true value $= 101.75$

$y_0 = 2$ $y_1 = 3$ $y_2 = 2$ $y_3 = 5$

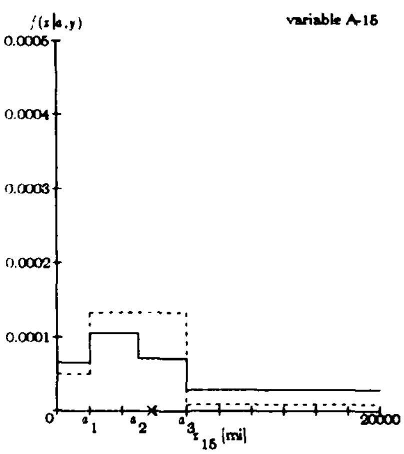
$a_1 = 10$ $a_2 = 15$ $a_3 = 25$



$n = 13$ true value $= 321.1$

$y_0 = 2$ $y_1 = 3$ $y_2 = 2$ $y_3 = 6$

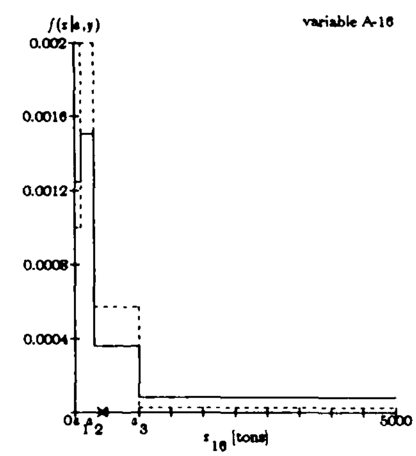
$a_1 = 150$ $a_2 = 350$ $a_3 = 700$



$n = 14$ true value $= 5864.25$

$y_0 = 2$ $y_1 = 4$ $y_2 = 2$ $y_3 = 6$

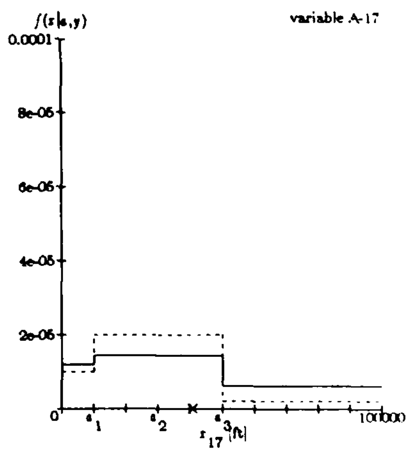
$a_1 = 2000$ $a_2 = 5000$ $a_3 = 8000$



$n = 15$ true value $= 425$

$y_0 = 2$ $y_1 = 4$ $y_2 = 3$ $y_3 = 6$

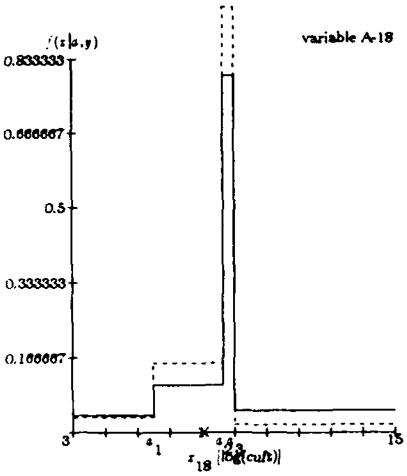
$a_1 = 100$ $a_2 = 300$ $a_3 = 1000$



$n = 16$ true value $= 40820$

$y_0 = 2$ $y_1 = 4$ $y_2 = 4$ $y_3 = 6$

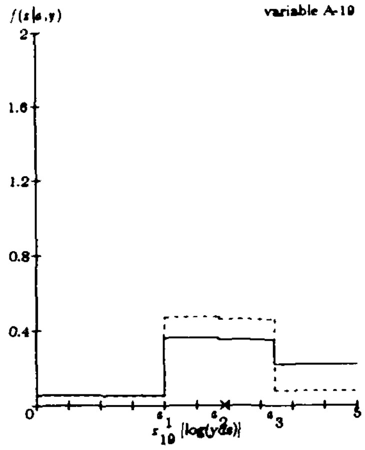
$a_1 = 10000$ $a_2 = 30000$ $a_3 = 50000$



$n = 17$ true value = 7.85

$y_0 = 2$ $y_1 = 4$ $y_2 = 5$ $y_3 = 6$

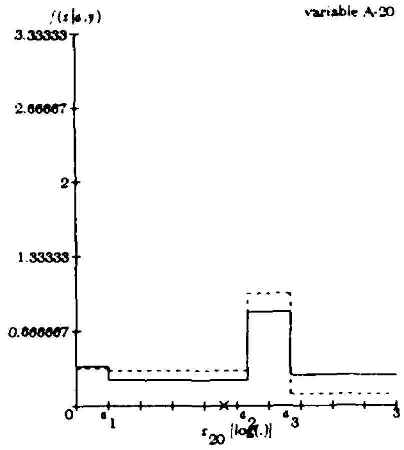
$a_1 = 6$ $a_2 = 8.6$ $a_3 = 9$



$n = 18$ true value = 2.94

$y_0 = 2$ $y_1 = 5$ $y_2 = 5$ $y_3 = 6$

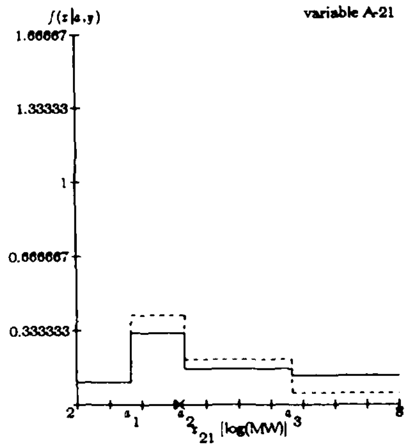
$a_1 = 2$ $a_2 = 2.85$ $a_3 = 3.72$



$n = 19$ true value = 1.38

$y_0 = 2$ $y_1 = 5$ $y_2 = 6$ $y_3 = 6$

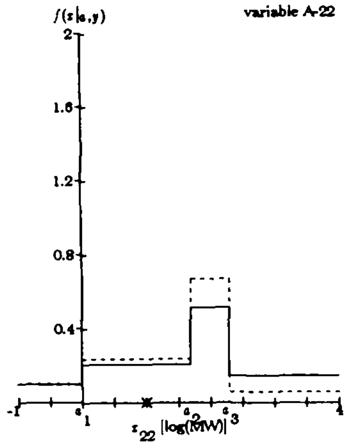
$a_1 = 0.3$ $a_2 = 1.6$ $a_3 = 2$



$n = 20$ true value = 3.897

$y_0 = 2$ $y_1 = 6$ $y_2 = 6$ $y_3 = 6$

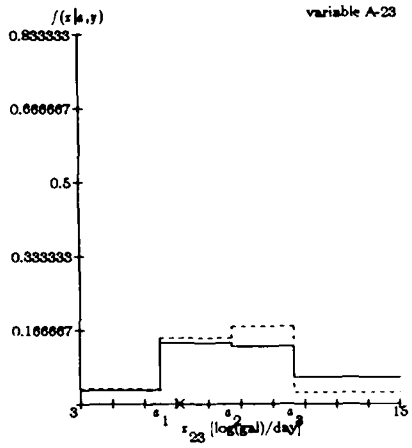
$a_1 = 3$ $a_2 = 4$ $a_3 = 6$



$n = 21$ true value = 1

$y_0 = 2$ $y_1 = 7$ $y_2 = 6$ $y_3 = 6$

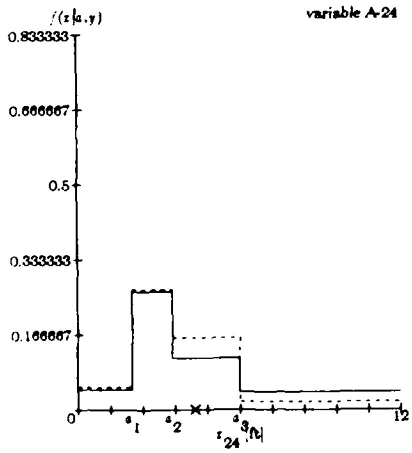
$a_1 = 0$ $a_2 = 1.7$ $a_3 = 2.3$



$n = 22$ true value = 6.7

$y_0 = 2$ $y_1 = 8$ $y_2 = 6$ $y_3 = 6$

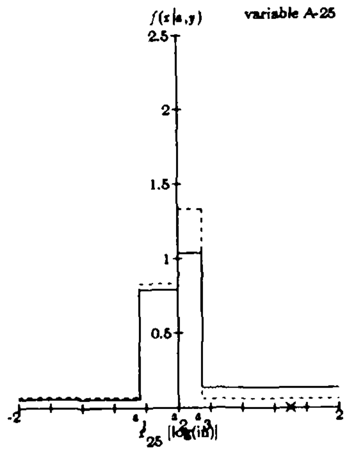
$a_1 = 6$ $a_2 = 8.7$ $a_3 = 11$



$n = 23$ true value = 4.33

$y_0 = 2$ $y_1 = 9$ $y_2 = 6$ $y_3 = 6$

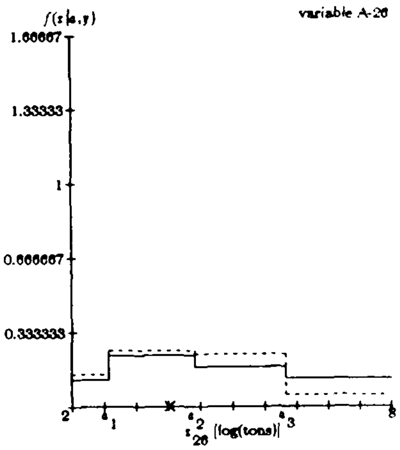
$a_1 = 2$ $a_2 = 3.5$ $a_3 = 6$



$n = 24$ true value = 1.398

$y_0 = 2$ $y_1 = 9$ $y_2 = 7$ $y_3 = 6$

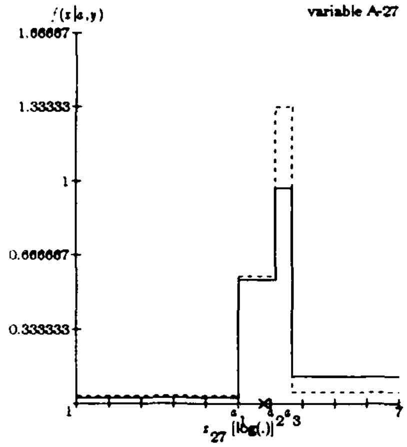
$a_1 = -0.48$ $a_2 = 0$ $a_3 = 0.3$



$n = 25$ true value = 3.82

$y_0 = 2$ $y_1 = 9$ $y_2 = 7$ $y_3 = 7$

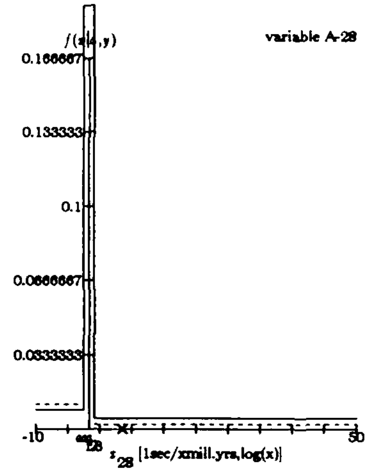
$a_1 = 2.7$ $a_2 = 4.3$ $a_3 = 6$



$n = 26$ true value $= 4.478$

$y_0 = 2$ $y_1 = 10$ $y_2 = 7$ $y_3 = 7$

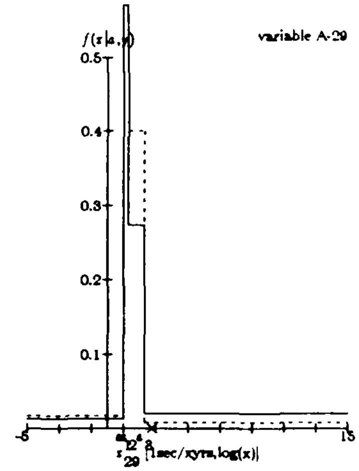
$a_1 = 4$ $a_2 = 4.7$ $a_3 = 5$



$n = 27$ true value $= 6.23$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 7$

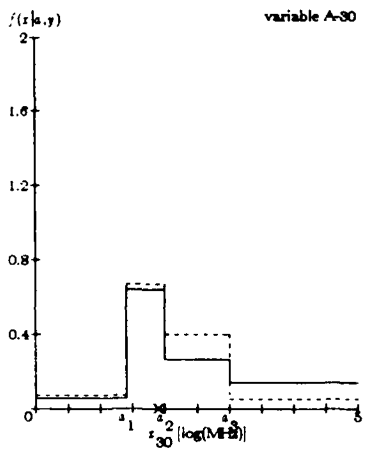
$a_1 = -1$ $a_2 = 0$ $a_3 = 1$



$n = 28$ true value $= 2.78$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 8$

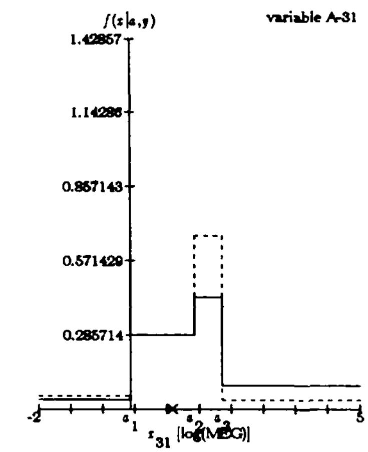
$a_1 = 1$ $a_2 = 1.3$ $a_3 = 2.3$



$n = 29$ true value $= 1.9$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 9$

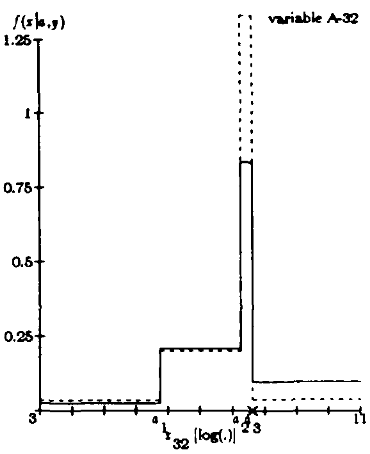
$a_1 = 1.4$ $a_2 = 2$ $a_3 = 3$



$n = 30$ true value $= 0.924$

$y_0 = 2$ $y_1 = 12$ $y_2 = 7$ $y_3 = 9$

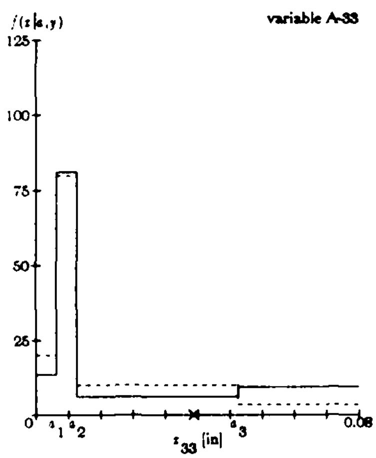
$a_1 = 0$ $a_2 = 1.4$ $a_3 = 2$



$n = 31$ true value $= 8.3$

$y_0 = 2$ $y_1 = 13$ $y_2 = 7$ $y_3 = 9$

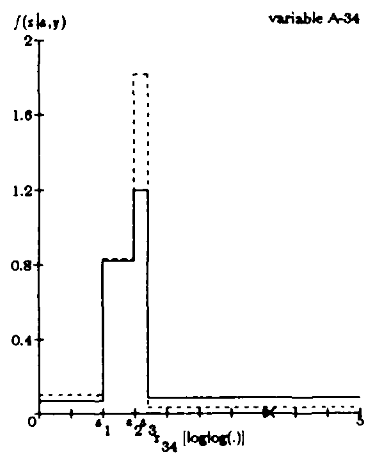
$a_1 = 6$ $a_2 = 8$ $a_3 = 8.3$



$n = 32$ true value $= 0.039$

$y_0 = 2$ $y_1 = 13$ $y_2 = 7$ $y_3 = 10$

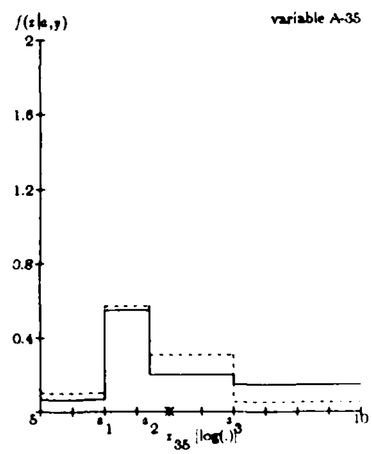
$a_1 = 0.005$ $a_2 = 0.01$ $a_3 = 0.05$



$n = 33$ true value $= 3.594$

$y_0 = 2$ $y_1 = 13$ $y_2 = 8$ $y_3 = 10$

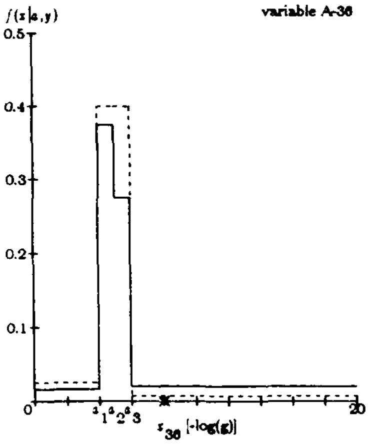
$a_1 = 1$ $a_2 = 1.48$ $a_3 = 1.7$



$n = 34$ true value $= 7.001$

$y_0 = 2$ $y_1 = 13$ $y_2 = 8$ $y_3 = 11$

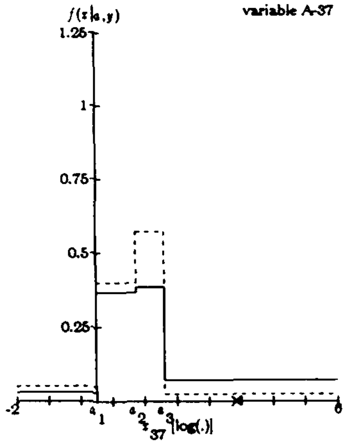
$a_1 = 6$ $a_2 = 6.7$ $a_3 = 8$



$n = 35$ true value = 8

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 11$

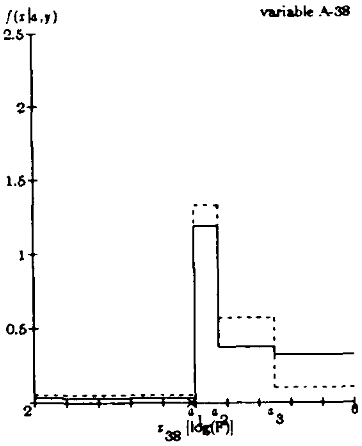
$a_1 = 4$ $a_2 = 5$ $a_3 = 6$



$n = 36$ true value = 3.477

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 12$

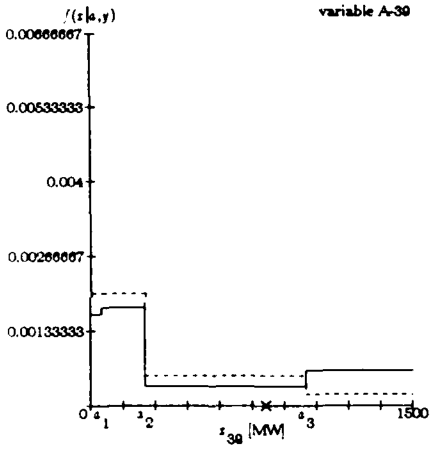
$a_1 = 0$ $a_2 = 1$ $a_3 = 1.7$



$n = 37$ true value = 3.955

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 13$

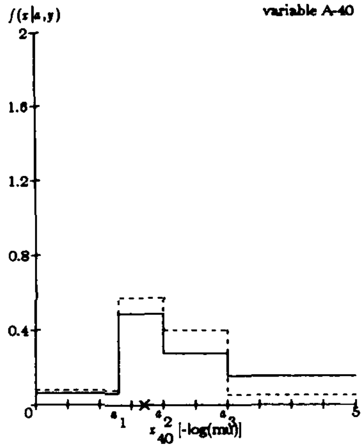
$a_1 = 4$ $a_2 = 4.3$ $a_3 = 5$



$n = 38$ true value = 815

$y_0 = 3$ $y_1 = 13$ $y_2 = 9$ $y_3 = 13$

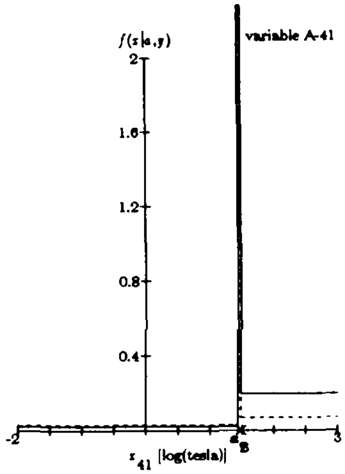
$a_1 = 50$ $a_2 = 250$ $a_3 = 1000$



$n = 39$ true value = 1.7

$y_0 = 3$ $y_1 = 13$ $y_2 = 10$ $y_3 = 13$

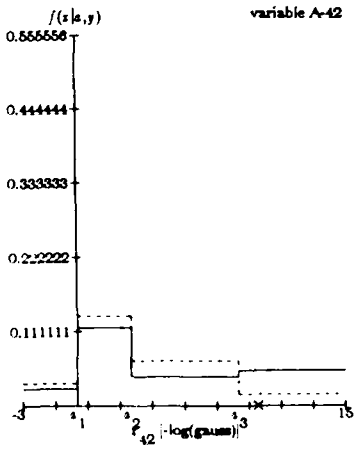
$a_1 = 1.3$ $a_2 = 2$ $a_3 = 3$



$n = 40$ true value = 1.48

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 13$

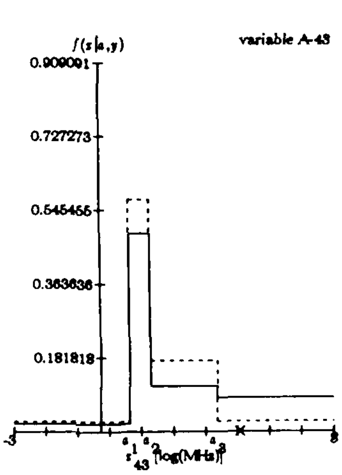
$a_1 = 1.43$ $a_2 = 1.45$ $a_3 = 1.48$



$n = 41$ true value = 10.1

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 14$

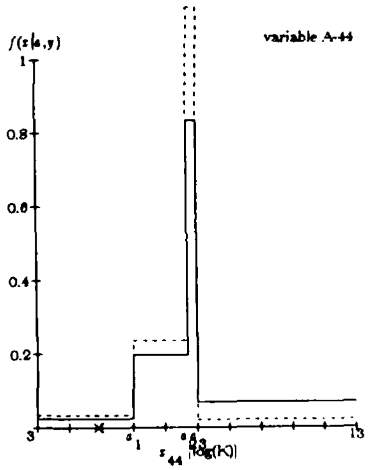
$a_1 = 0$ $a_2 = 3$ $a_3 = 9$



$n = 42$ true value = 4.778

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 15$

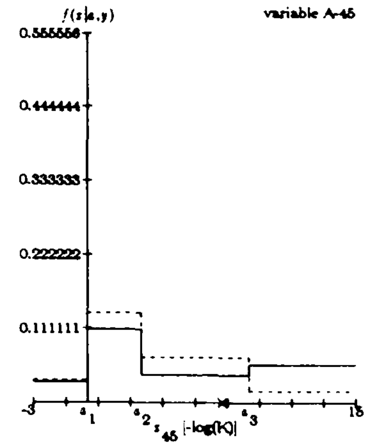
$a_1 = 1$ $a_2 = 1.7$ $a_3 = 4$



$n = 43$ true value = 4.914

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 16$

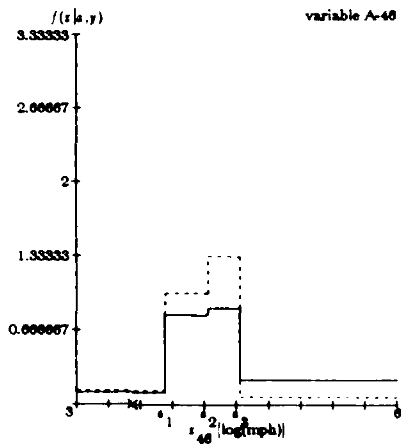
$a_1 = 6$ $a_2 = 7.7$ $a_3 = 8$



$n = 44$ true value $= 7.52$

$y_0 = 4$ $y_1 = 14$ $y_2 = 10$ $y_3 = 16$

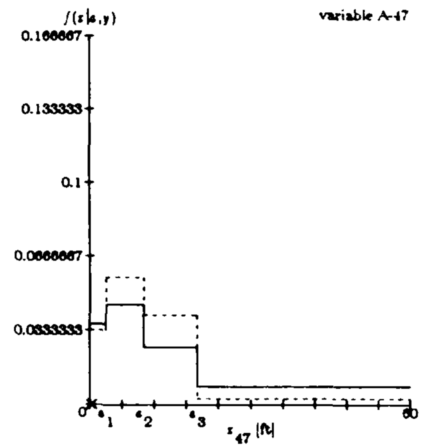
$a_1 = 0$ $a_2 = 3$ $a_3 = 9$



$n = 45$ true value $= 3.526$

$y_0 = 4$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

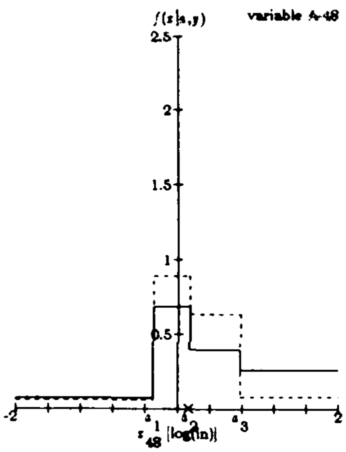
$a_1 = 3.88$ $a_2 = 4.23$ $a_3 = 4.53$



$n = 46$ true value $= 0.42$

$y_0 = 5$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

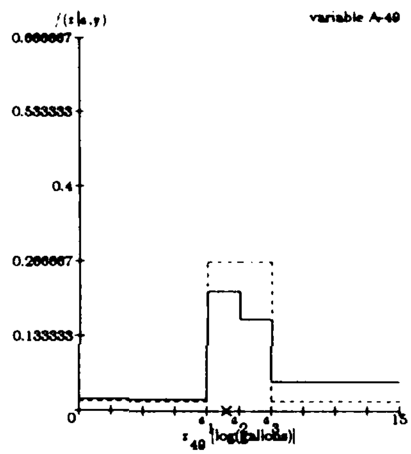
$a_1 = 3$ $a_2 = 10$ $a_3 = 20$



$n = 47$ true value $= 0.138$

$y_0 = 6$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

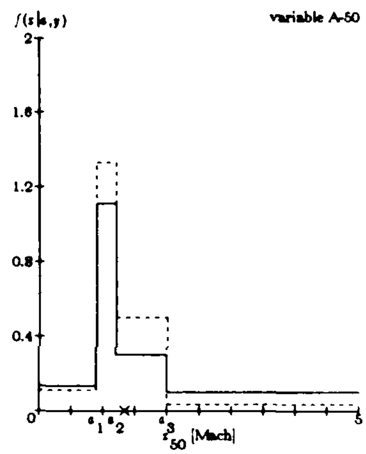
$a_1 = -0.3$ $a_2 = 0.15$ $a_3 = 0.78$



$n = 48$ true value $= 6.92$

$y_0 = 6$ $y_1 = 15$ $y_2 = 11$ $y_3 = 16$

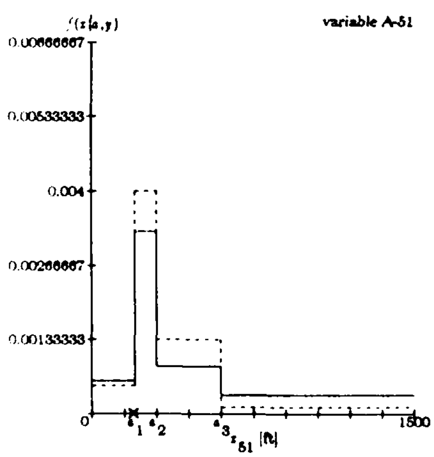
$a_1 = 6$ $a_2 = 7.5$ $a_3 = 9$



$n = 49$ true value $= 1.34$

$y_0 = 6$ $y_1 = 16$ $y_2 = 11$ $y_3 = 16$

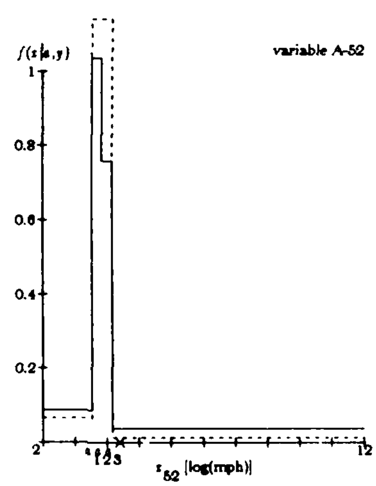
$a_1 = 0.9$ $a_2 = 1.2$ $a_3 = 2$



$n = 50$ true value $= 191.94$

$y_0 = 6$ $y_1 = 16$ $y_2 = 12$ $y_3 = 16$

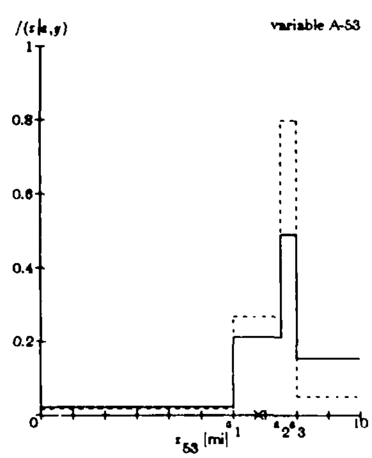
$a_1 = 200$ $a_2 = 300$ $a_3 = 600$



$n = 51$ true value $= 4.394$

$y_0 = 7$ $y_1 = 16$ $y_2 = 12$ $y_3 = 16$

$a_1 = 3.54$ $a_2 = 3.86$ $a_3 = 4.18$



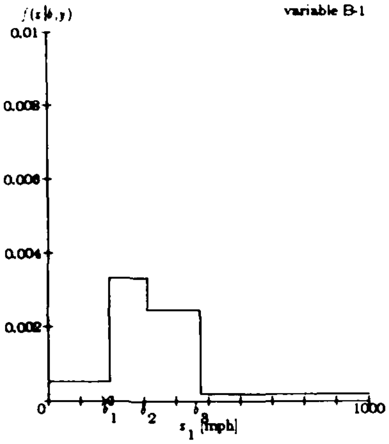
$n = 52$ true value $= 6.78$

$y_0 = 7$ $y_1 = 16$ $y_2 = 12$ $y_3 = 17$

$a_1 = 6$ $a_2 = 7.5$ $a_3 = 8$

11.2 Appendix 3B

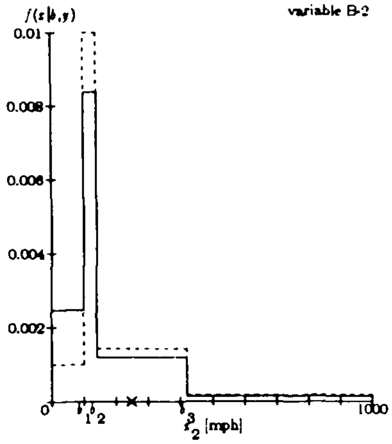
Advice (dashed) and corrected advice (solid) of source B



$n = 0$ true value $= 179.9$

$y_0 = 0$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

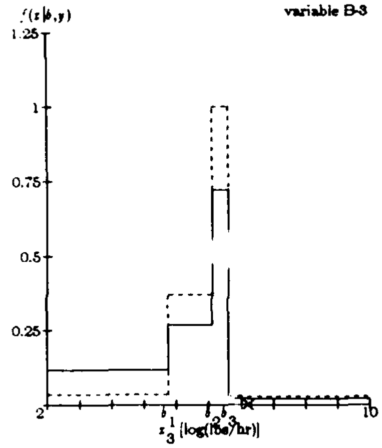
$b_1 = 192$ $b_2 = 312$ $b_3 = 476$



$n = 1$ true value $= 247.93$

$y_0 = 1$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

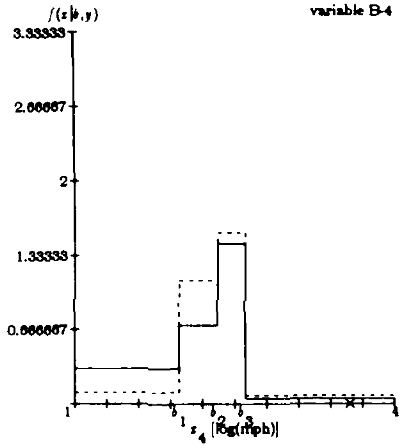
$b_1 = 100$ $b_2 = 140$ $b_3 = 420$



$n = 2$ true value $= 6.97$

$y_0 = 2$ $y_1 = 0$ $y_2 = 0$ $y_3 = 0$

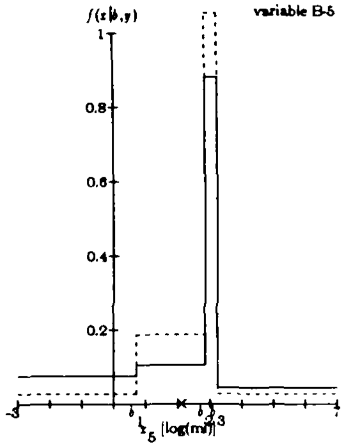
$b_1 = 5$ $b_2 = 6.08$ $b_3 = 6.48$



$n = 3$ true value $= 3.58$

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 0$

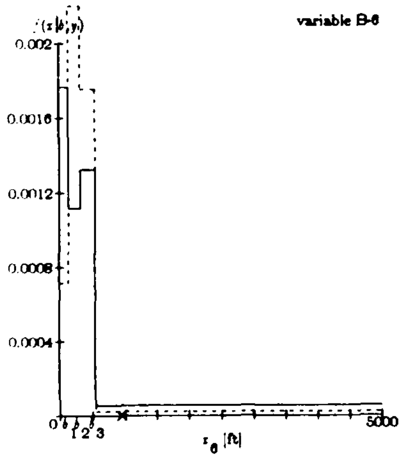
$b_1 = 1.98$ $b_2 = 2.34$ $b_3 = 2.6$



$n = 4$ true value $= 2.114$

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 1$

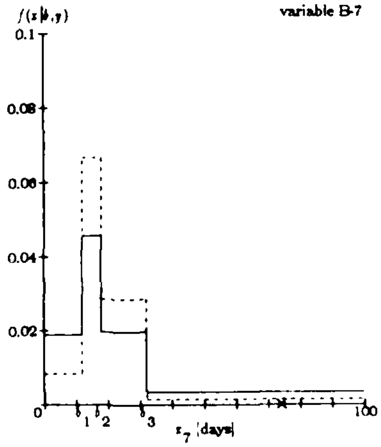
$b_1 = 0.7$ $b_2 = 2.85$ $b_3 = 3.23$



$n = 5$ true value $= 950$

$y_0 = 2$ $y_1 = 0$ $y_2 = 1$ $y_3 = 2$

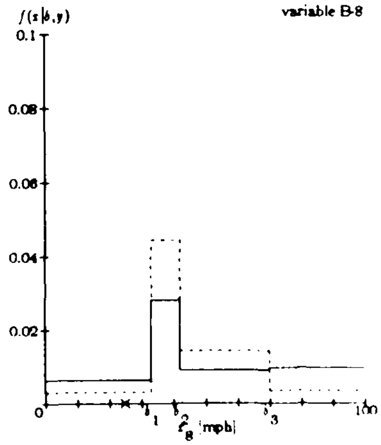
$b_1 = 140$ $b_2 = 322$ $b_3 = 550$



$n = 6$ true value $= 74$

$y_0 = 2$ $y_1 = 1$ $y_2 = 1$ $y_3 = 2$

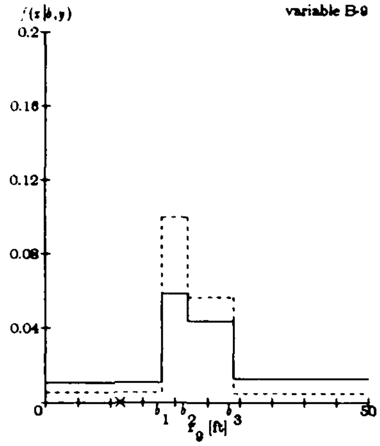
$b_1 = 12$ $b_2 = 18$ $b_3 = 32$



$n = 7$ true value $= 24.74$

$y_0 = 2$ $y_1 = 1$ $y_2 = 1$ $y_3 = 3$

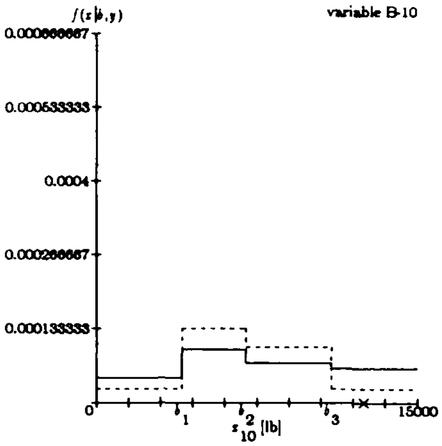
$b_1 = 33$ $b_2 = 42$ $b_3 = 70$



$n = 8$ true value = 11.5

$y_0 = 2$ $y_1 = 1$ $y_2 = 2$ $y_3 = 3$

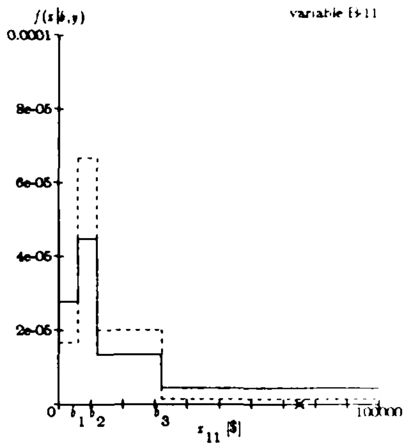
$b_1 = 18$ $b_2 = 22$ $b_3 = 29.1$



$n = 9$ true value = 12500

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 3$

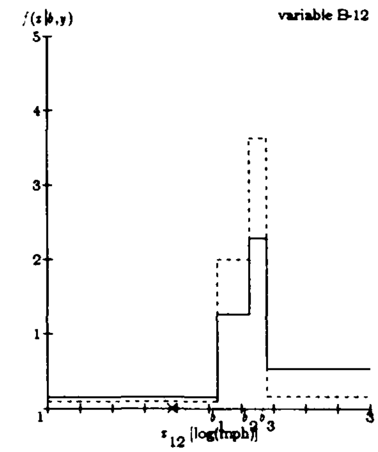
$b_1 = 4000$ $b_2 = 7000$ $b_3 = 11000$



$n = 10$ true value = 75000

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 4$

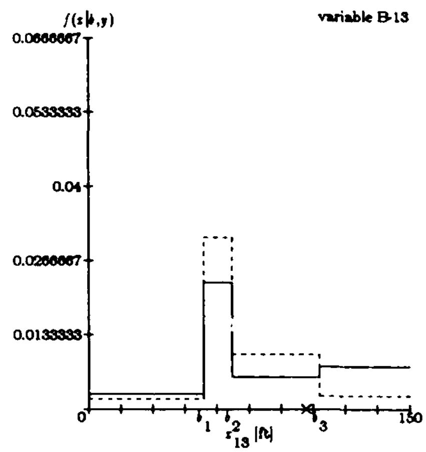
$b_1 = 6000$ $b_2 = 12000$ $b_3 = 30000$



$n = 11$ true value = 1.77

$y_0 = 2$ $y_1 = 2$ $y_2 = 2$ $y_3 = 5$

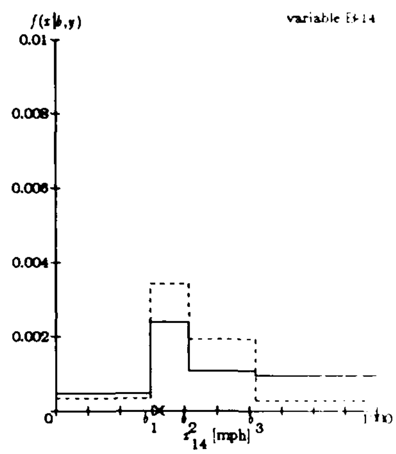
$b_1 = 2.05$ $b_2 = 2.25$ $b_3 = 2.36$



$n = 12$ true value = 101.75

$y_0 = 2$ $y_1 = 3$ $y_2 = 2$ $y_3 = 5$

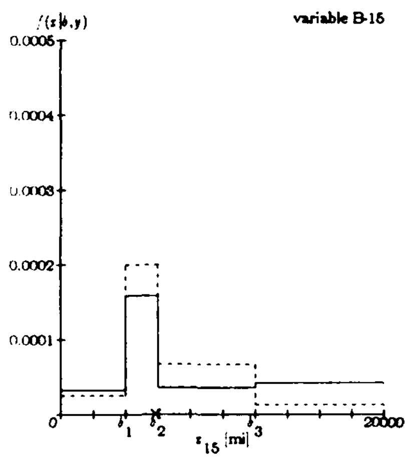
$b_1 = 54$ $b_2 = 67$ $b_3 = 108$



$n = 13$ true value = 321.1

$y_0 = 2$ $y_1 = 3$ $y_2 = 2$ $y_3 = 6$

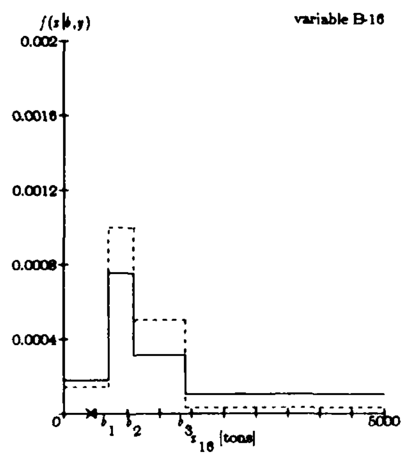
$b_1 = 295$ $b_2 = 412$ $b_3 = 620$



$n = 14$ true value = 5864.25

$y_0 = 2$ $y_1 = 4$ $y_2 = 2$ $y_3 = 6$

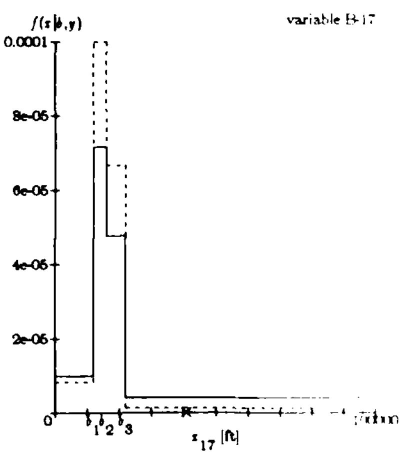
$b_1 = 4000$ $b_2 = 6000$ $b_3 = 12000$



$n = 15$ true value = 425

$y_0 = 2$ $y_1 = 4$ $y_2 = 3$ $y_3 = 6$

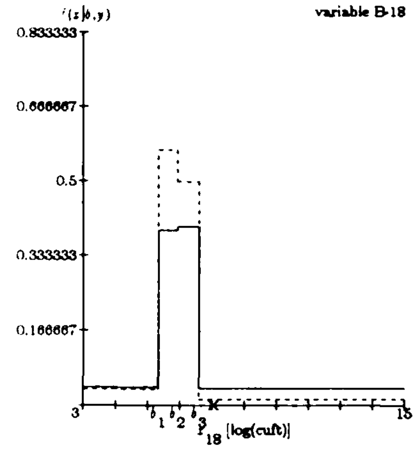
$b_1 = 700$ $b_2 = 1100$ $b_3 = 1900$



$n = 16$ true value = 40820

$y_0 = 2$ $y_1 = 4$ $y_2 = 4$ $y_3 = 6$

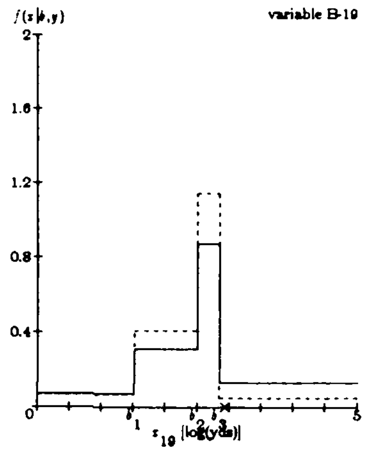
$b_1 = 12000$ $b_2 = 16000$ $b_3 = 22000$



$n = 17$ true value = 7.85

$y_0 = 2$ $y_1 = 4$ $y_2 = 5$ $y_3 = 6$

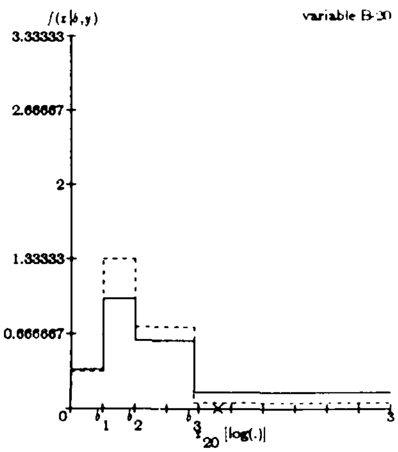
$b_1 = 5.8$ $b_2 = 6.5$ $b_3 = 7.3$



$n = 18$ true value = 2.94

$y_0 = 2$ $y_1 = 5$ $y_2 = 5$ $y_3 = 6$

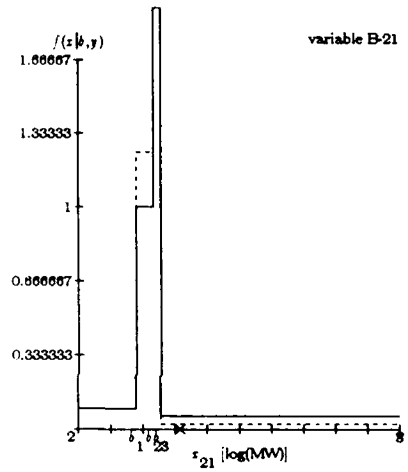
$b_1 = 1.52$ $b_2 = 2.5$ $b_3 = 2.85$



$n = 19$ true value = 1.38

$y_0 = 2$ $y_1 = 5$ $y_2 = 6$ $y_3 = 6$

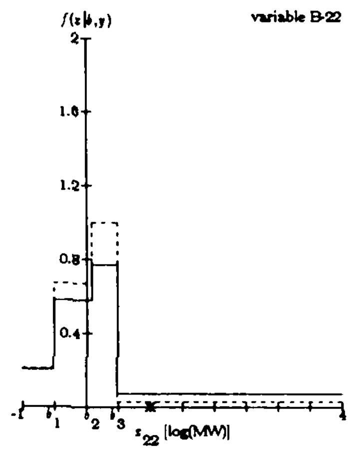
$b_1 = 0.3$ $b_2 = 0.6$ $b_3 = 1.15$



$n = 20$ true value = 3.897

$y_0 = 2$ $y_1 = 6$ $y_2 = 6$ $y_3 = 6$

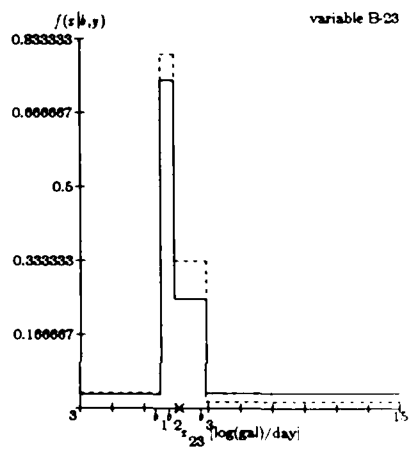
$b_1 = 3.08$ $b_2 = 3.4$ $b_3 = 3.54$



$n = 21$ true value = 1

$y_0 = 2$ $y_1 = 7$ $y_2 = 6$ $y_3 = 6$

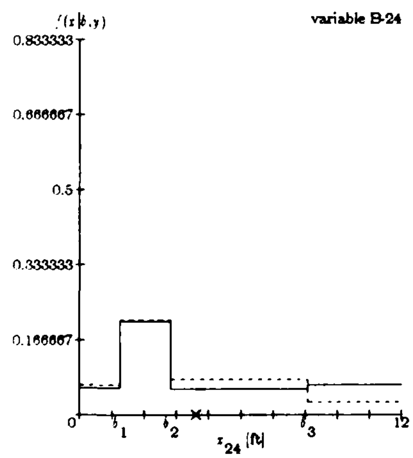
$b_1 = -0.52$ $b_2 = 0.08$ $b_3 = 0.48$



$n = 22$ true value = 6.7

$y_0 = 2$ $y_1 = 8$ $y_2 = 6$ $y_3 = 6$

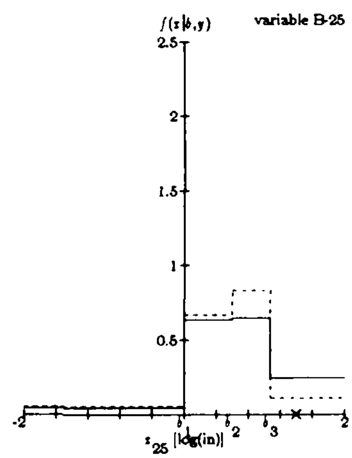
$b_1 = 6$ $b_2 = 6.5$ $b_3 = 7.7$



$n = 23$ true value = 4.33

$y_0 = 2$ $y_1 = 9$ $y_2 = 6$ $y_3 = 6$

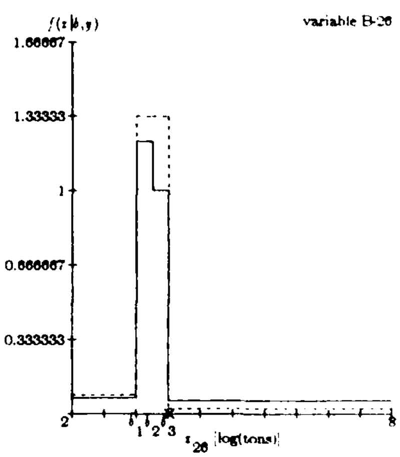
$b_1 = 1.5$ $b_2 = 3.4$ $b_3 = 8.5$



$n = 24$ true value = 1.398

$y_0 = 2$ $y_1 = 9$ $y_2 = 7$ $y_3 = 6$

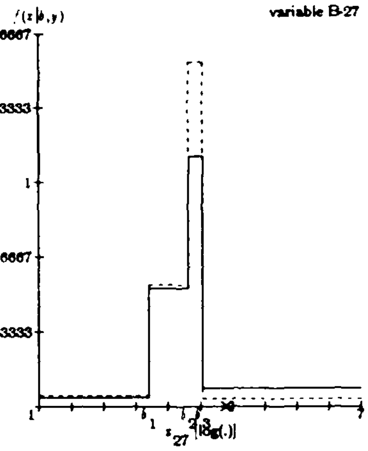
$b_1 = 0$ $b_2 = 0.6$ $b_3 = 1.08$



$n = 25$ true value = 3.82

$y_0 = 2$ $y_1 = 9$ $y_2 = 7$ $y_3 = 7$

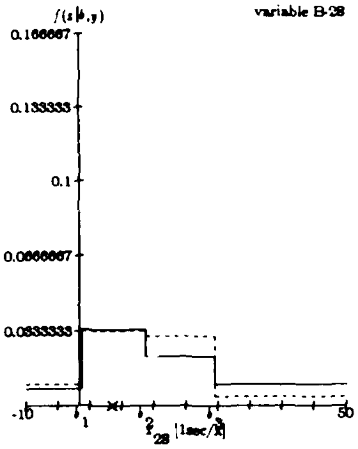
$b_1 = 3.2$ $b_2 = 3.5$ $b_3 = 3.8$



$n = 26$ true value $= 4.478$

$y_0 = 2$ $y_1 = 10$ $y_2 = 7$ $y_3 = 7$

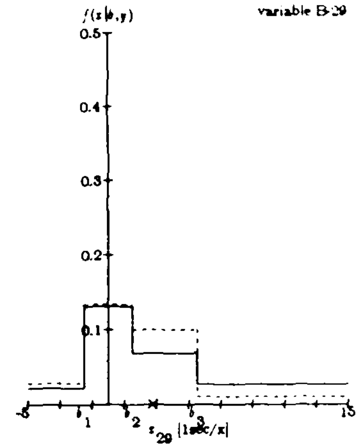
$b_1 = 3.04$ $b_2 = 3.78$ $b_3 = 4.04$



$n = 27$ true value $= 6.23$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 7$

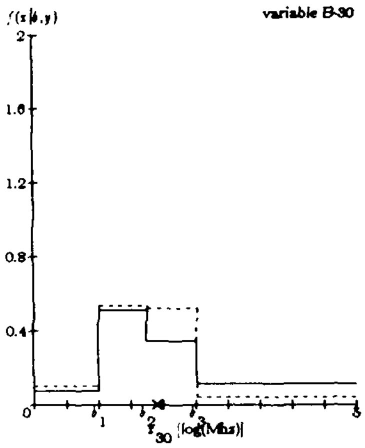
$b_1 = 0.5$ $b_2 = 12.5$ $b_3 = 25.5$



$n = 28$ true value $= 2.78$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 8$

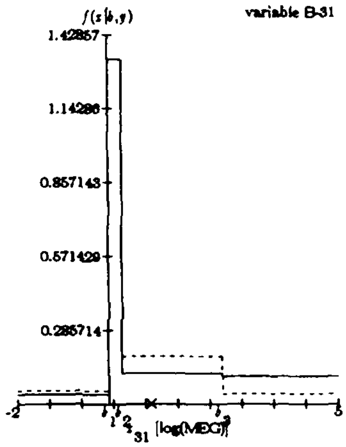
$b_1 = -1.5$ $b_2 = 1.5$ $b_3 = 5.5$



$n = 29$ true value $= 1.9$

$y_0 = 2$ $y_1 = 11$ $y_2 = 7$ $y_3 = 9$

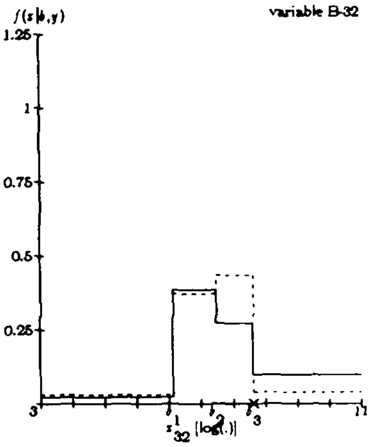
$b_1 = 1$ $b_2 = 1.75$ $b_3 = 2.52$



$n = 30$ true value $= 0.924$

$y_0 = 2$ $y_1 = 12$ $y_2 = 7$ $y_3 = 9$

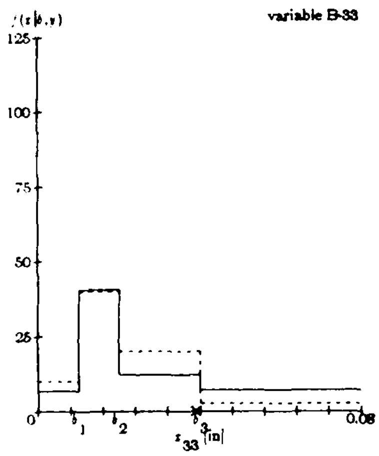
$b_1 = 0$ $b_2 = 0.3$ $b_3 = 2.48$



$n = 31$ true value $= 8.3$

$y_0 = 2$ $y_1 = 13$ $y_2 = 7$ $y_3 = 9$

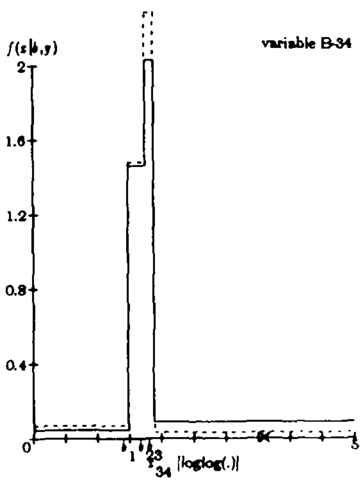
$b_1 = 6.3$ $b_2 = 7.38$ $b_3 = 8.3$



$n = 32$ true value $= 0.039$

$y_0 = 2$ $y_1 = 13$ $y_2 = 7$ $y_3 = 10$

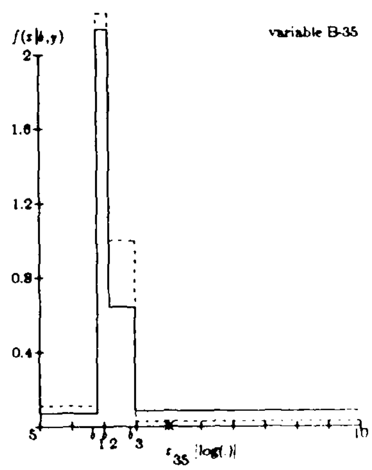
$b_1 = 0.01$ $b_2 = 0.02$ $b_3 = 0.04$



$n = 33$ true value $= 3.594$

$y_0 = 2$ $y_1 = 13$ $y_2 = 8$ $y_3 = 10$

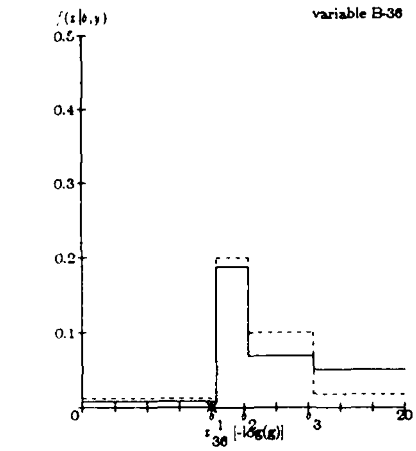
$b_1 = 1.48$ $b_2 = 1.75$ $b_3 = 1.88$



$n = 34$ true value $= 7.001$

$y_0 = 2$ $y_1 = 13$ $y_2 = 8$ $y_3 = 11$

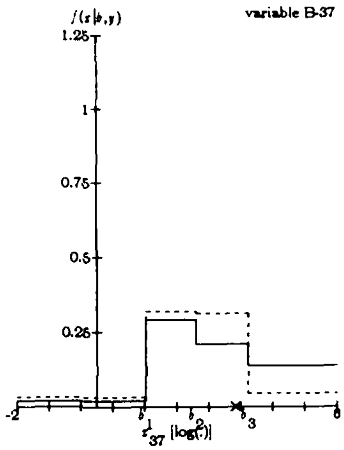
$b_1 = 5.9$ $b_2 = 6.08$ $b_3 = 6.48$



$n = 35$ true value = 8

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 11$

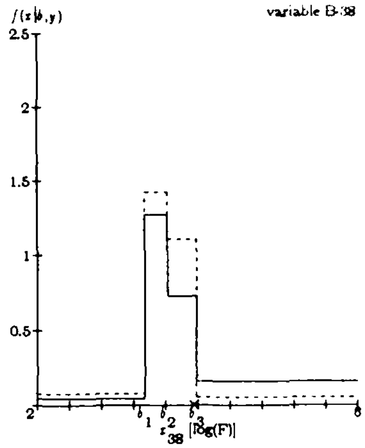
$b_1 = 8.3$ $b_2 = 10.3$ $b_3 = 14.3$



$n = 36$ true value = 3.477

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 12$

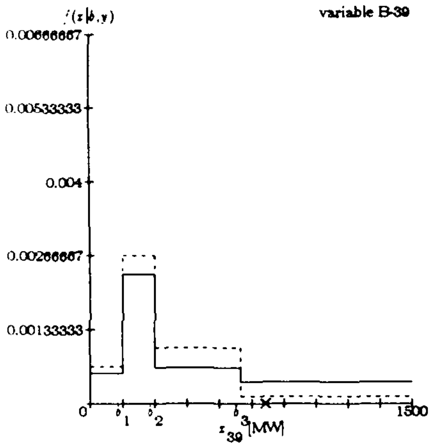
$b_1 = 1.23$ $b_2 = 2.49$ $b_3 = 3.78$



$n = 37$ true value = 3.955

$y_0 = 2$ $y_1 = 13$ $y_2 = 9$ $y_3 = 13$

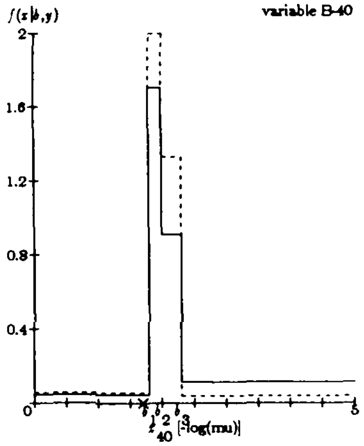
$b_1 = 3.34$ $b_2 = 3.62$ $b_3 = 3.98$



$n = 38$ true value = 815

$y_0 = 3$ $y_1 = 13$ $y_2 = 9$ $y_3 = 13$

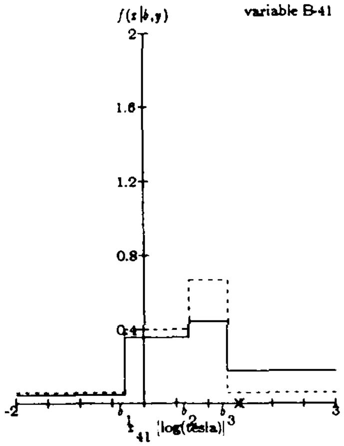
$b_1 = 150$ $b_2 = 300$ $b_3 = 700$



$n = 39$ true value = 1.7

$y_0 = 3$ $y_1 = 13$ $y_2 = 10$ $y_3 = 13$

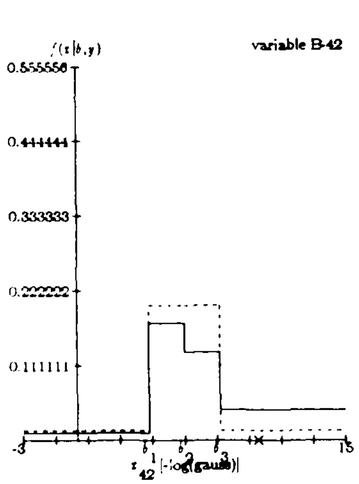
$b_1 = 1.8$ $b_2 = 2$ $b_3 = 2.3$



$n = 40$ true value = 1.48

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 13$

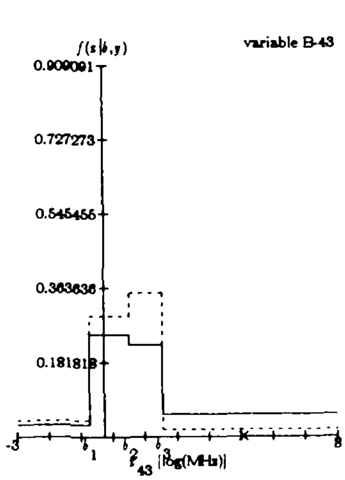
$b_1 = -0.3$ $b_2 = 0.7$ $b_3 = 1.3$



$n = 41$ true value = 10.1

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 14$

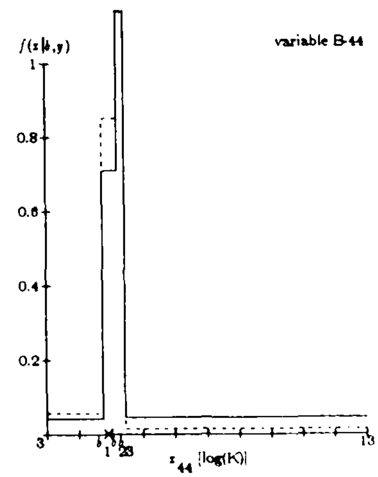
$b_1 = 4$ $b_2 = 6$ $b_3 = 8$



$n = 42$ true value = 4.778

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 15$

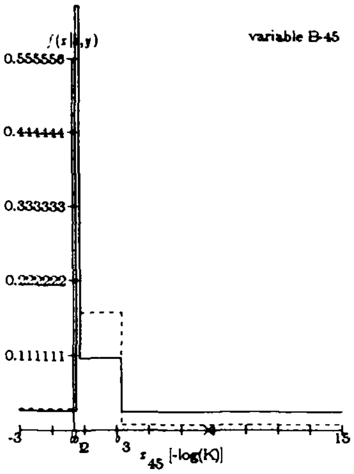
$b_1 = -0.5$ $b_2 = 0.86$ $b_3 = 2$



$n = 43$ true value = 4.914

$y_0 = 3$ $y_1 = 14$ $y_2 = 10$ $y_3 = 16$

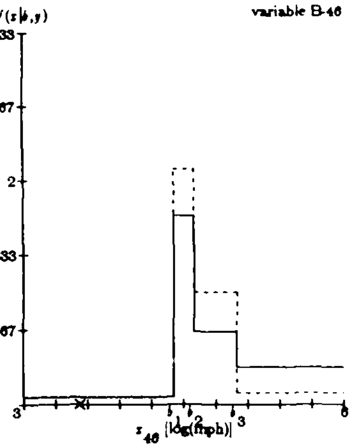
$b_1 = 4.75$ $b_2 = 5.22$ $b_3 = 5.44$



$n = 44$ true value = 7.52

$y_0 = 4$ $y_1 = 14$ $y_2 = 10$ $y_3 = 16$

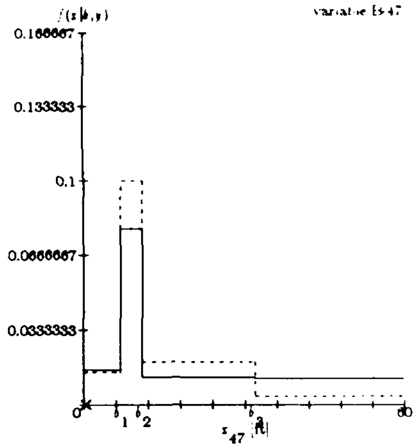
$b_1 = 0.2$ $b_2 = 0.4$ $b_3 = 2.7$



$n = 45$ true value = 3.525

$y_0 = 4$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

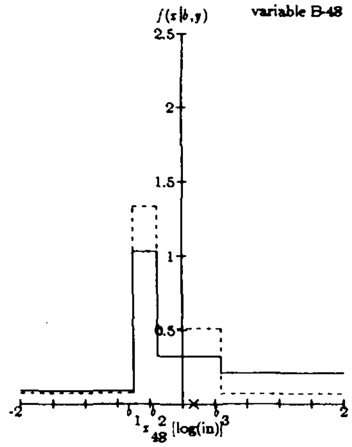
$b_1 = 4.41$ $b_2 = 4.6$ $b_3 = 5$



$n = 46$ true value = 0.42

$y_0 = 5$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

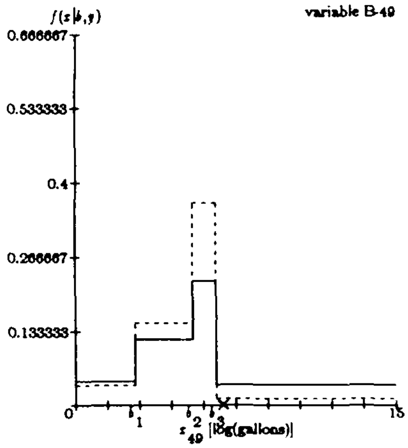
$b_1 = 7$ $b_2 = 11$ $b_3 = 32$



$n = 47$ true value = 0.138

$y_0 = 6$ $y_1 = 14$ $y_2 = 11$ $y_3 = 16$

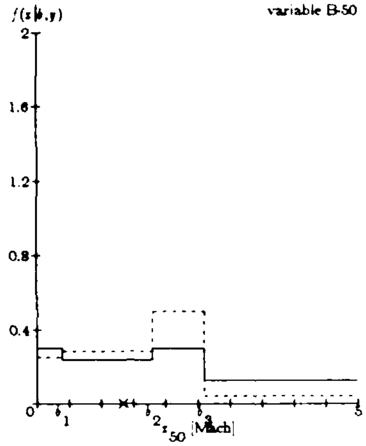
$b_1 = -0.6$ $b_2 = -0.3$ $b_3 = 0.48$



$n = 48$ true value = 6.92

$y_0 = 6$ $y_1 = 15$ $y_2 = 11$ $y_3 = 16$

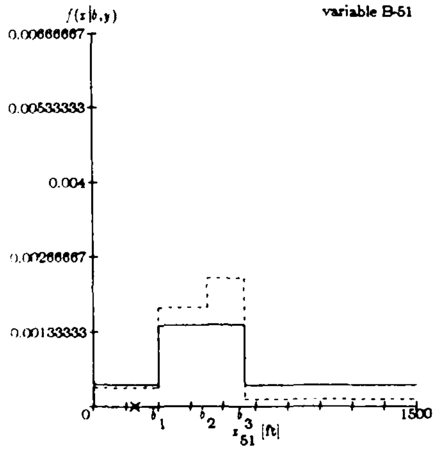
$b_1 = 2.8$ $b_2 = 5.5$ $b_3 = 6.6$



$n = 49$ true value = 1.34

$y_0 = 6$ $y_1 = 16$ $y_2 = 11$ $y_3 = 16$

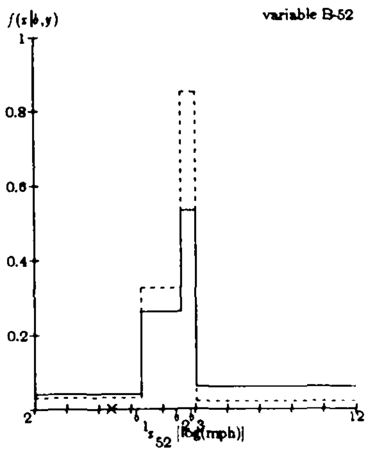
$b_1 = 0.4$ $b_2 = 1.8$ $b_3 = 2.6$



$n = 50$ true value = 191.94

$y_0 = 6$ $y_1 = 16$ $y_2 = 12$ $y_3 = 16$

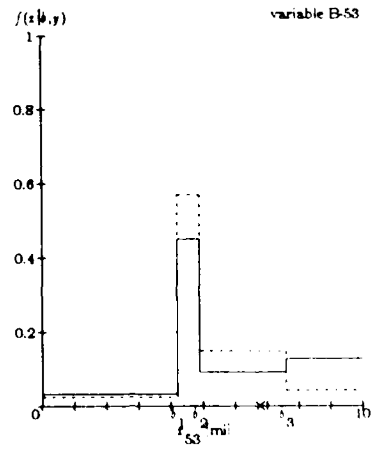
$b_1 = 300$ $b_2 = 526$ $b_3 = 700$



$n = 51$ true value = 4.394

$y_0 = 7$ $y_1 = 16$ $y_2 = 12$ $y_3 = 16$

$b_1 = 5.33$ $b_2 = 6.56$ $b_3 = 7.03$



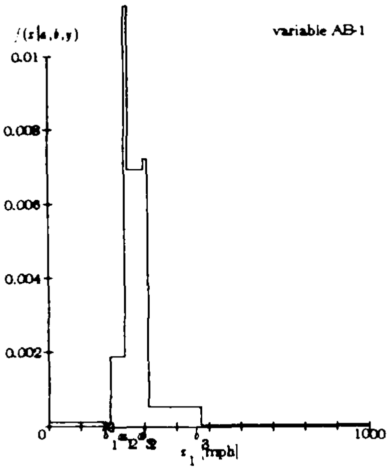
$n = 52$ true value = 6.78

$y_0 = 7$ $y_1 = 16$ $y_2 = 12$ $y_3 = 17$

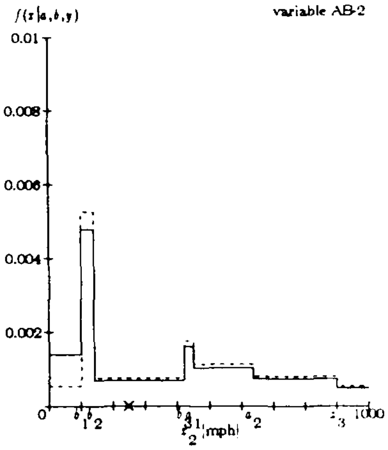
$b_1 = 4.2$ $b_2 = 4.9$ $b_3 = 7.6$

11.3 Appendix 3AB

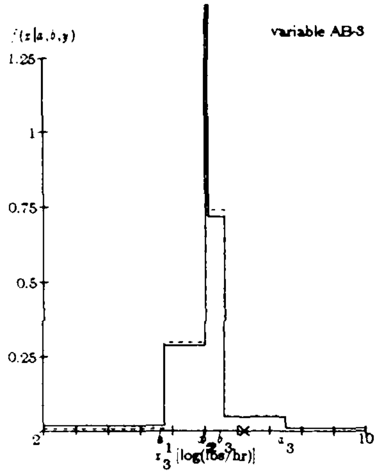
Advice (dashed) and corrected advice (solid) of sources A and B together



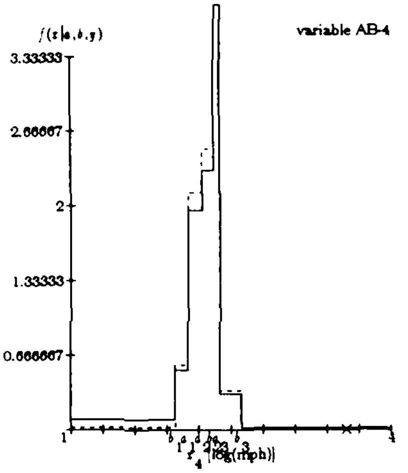
$n = 0$ true value $= 179.9$
$y_{00} = 0$ $y_{01} = 0$ $y_{11} = 0$ $y_{21} = 0$
$y_{31} = 0$ $y_{32} = 0$ $y_{33} = 0$
$b_1 = 192$ $b_2 = 312$ $b_3 = 476$
$a_1 = 238$ $a_2 = 250$ $a_3 = 300$



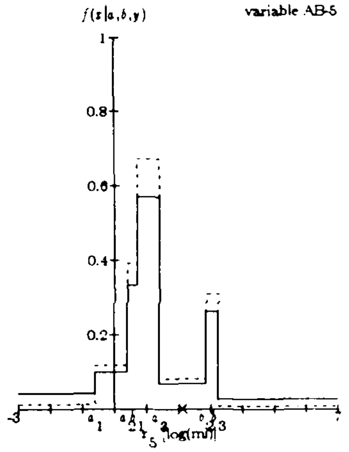
$n = 1$ true value $= 247.93$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 0$ $y_{03} = 0$
$y_{13} = 0$ $y_{23} = 0$ $y_{33} = 0$
$b_1 = 100$ $b_2 = 140$ $b_3 = 420$
$a_1 = 450$ $a_2 = 636$ $a_3 = 900$



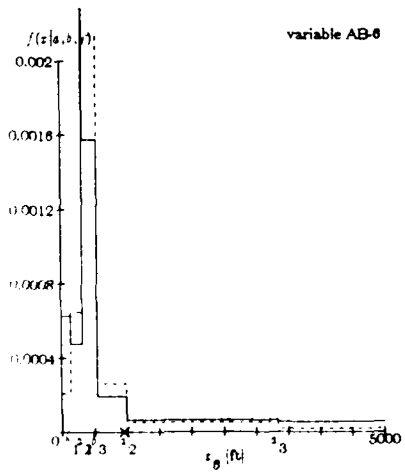
$n = 1$ true value $= 6.97$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 0$ $y_{21} = 0$
$y_{22} = 0$ $y_{23} = 0$ $y_{33} = 0$
$b_1 = 5$ $b_2 = 6.08$ $b_3 = 6.48$
$a_1 = 5$ $a_2 = 6$ $a_3 = 8$



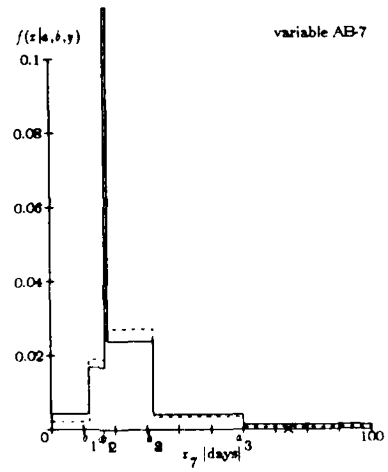
$n = 1$ true value $= 3.58$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 0$ $y_{21} = 0$
$y_{22} = 0$ $y_{32} = 0$ $y_{33} = 0$
$b_1 = 1.98$ $b_2 = 2.34$ $b_3 = 2.6$
$a_1 = 2.1$ $a_2 = 2.23$ $a_3 = 2.4$



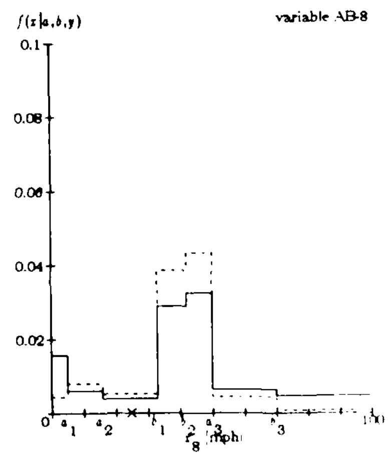
$n = 2$ true value $= 2.114$
$y_{00} = 1$ $y_{10} = 0$ $y_{20} = 0$ $y_{21} = 0$
$y_{31} = 0$ $y_{32} = 0$ $y_{33} = 1$
$b_1 = 0.7$ $b_2 = 2.85$ $b_3 = 3.23$
$a_1 = -0.6$ $a_2 = 0.4$ $a_3 = 1.4$



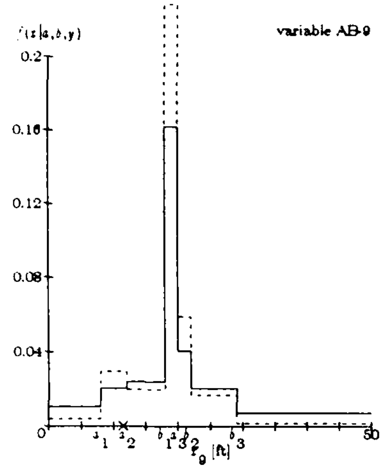
$n = 4$ true value $= 950$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 1$ $y_{12} = 0$
$y_{13} = 0$ $y_{23} = 1$ $y_{33} = 1$
$b_1 = 140$ $b_2 = 322$ $b_3 = 550$
$a_1 = 330$ $a_2 = 1000$ $a_3 = 3340$



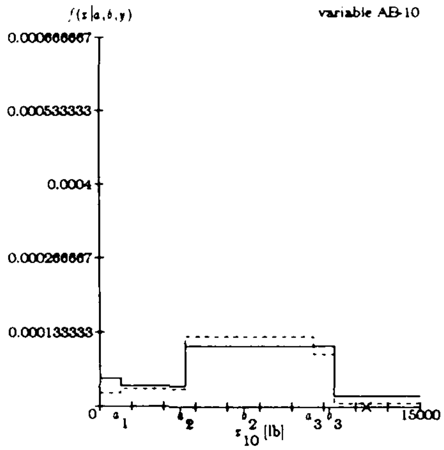
$n = 4$ true value $= 74$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 0$ $y_{12} = 0$
$y_{13} = 1$ $y_{23} = 1$ $y_{33} = 1$
$b_1 = 12$ $b_2 = 18$ $b_3 = 32$
$a_1 = 17$ $a_2 = 32$ $a_3 = 60$



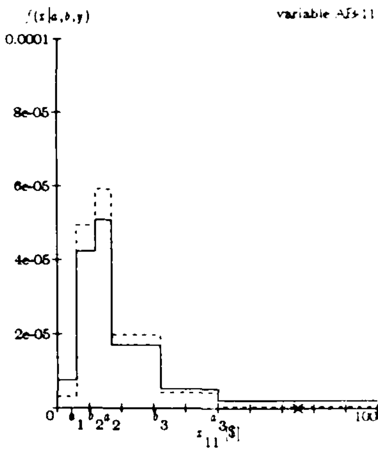
$n = 4$ true value $= 24.74$
$y_{00} = 1$ $y_{10} = 0$ $y_{20} = 0$ $y_{21} = 0$
$y_{22} = 0$ $y_{32} = 1$ $y_{33} = 2$
$b_1 = 33$ $b_2 = 42$ $b_3 = 70$
$a_1 = 5$ $a_2 = 16$ $a_3 = 50$



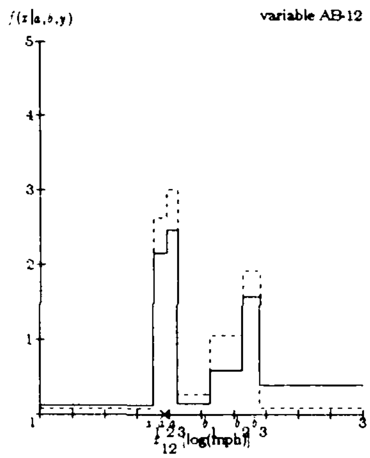
$n = 5$ true value $= 11.5$
$y_{00} = 1$ $y_{10} = 0$ $y_{20} = 1$ $y_{21} = 0$
$y_{31} = 0$ $y_{32} = 1$ $y_{33} = 2$
$b_1 = 18$ $b_2 = 22$ $b_3 = 29.1$
$a_1 = 8$ $a_2 = 12$ $a_3 = 20$



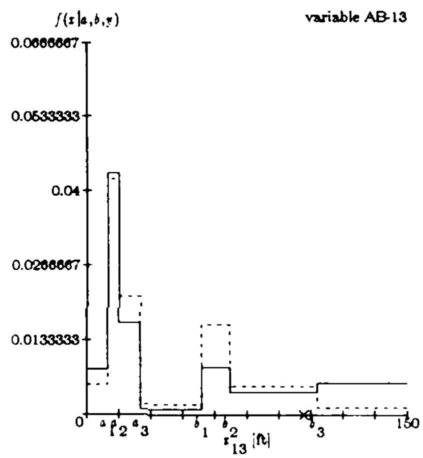
$n = 5$ true value $= 12500$
$y_{00} = 1$ $y_{10} = 1$ $y_{11} = 0$ $y_{21} = 0$
$y_{22} = 0$ $y_{32} = 1$ $y_{33} = 2$
$b_1 = 4000$ $b_2 = 7000$ $b_3 = 11000$
$a_1 = 1000$ $a_2 = 4000$ $a_3 = 10000$



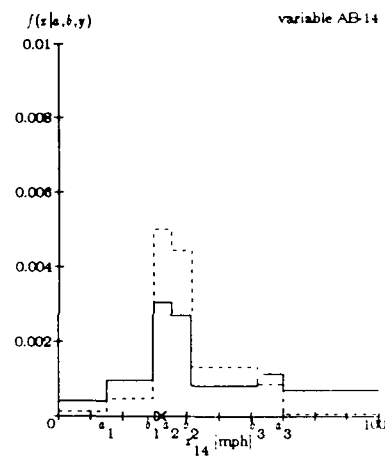
$n = 5$ true value $= 75000$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 0$ $y_{12} = 0$
$y_{22} = 0$ $y_{23} = 1$ $y_{33} = 3$
$b_1 = 6000$ $b_2 = 12000$ $b_3 = 32000$
$a_1 = 6000$ $a_2 = 17000$ $a_3 = 50000$



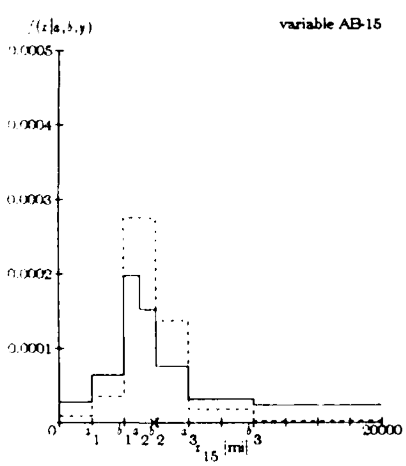
$n = 8$ true value $= 1.77$
$y_{00} = 1$ $y_{10} = 1$ $y_{20} = 1$ $y_{30} = 0$
$y_{31} = 0$ $y_{32} = 1$ $y_{33} = 4$
$b_1 = 2.05$ $b_2 = 2.25$ $b_3 = 2.36$
$a_1 = 1.7$ $a_2 = 1.78$ $a_3 = 1.85$



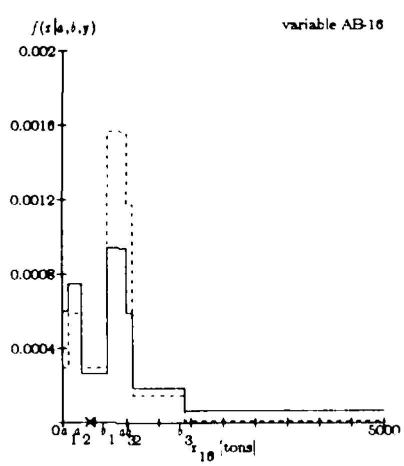
$n = 9$ true value $= 101.75$
$y_{00} = 1$ $y_{10} = 2$ $y_{20} = 1$ $y_{30} = 0$
$y_{31} = 0$ $y_{32} = 1$ $y_{33} = 4$
$b_1 = 54$ $b_2 = 67$ $b_3 = 108$
$a_1 = 10$ $a_2 = 15$ $a_3 = 25$



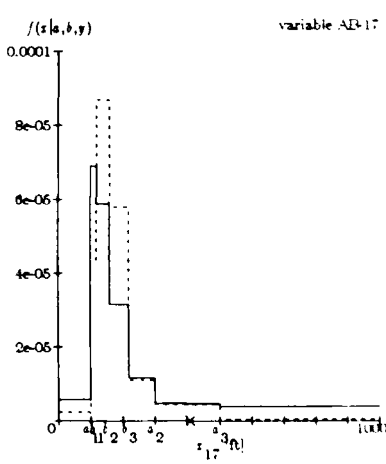
$n = 8$ true value $= 321.1$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 0$ $y_{21} = 0$
$y_{22} = 0$ $y_{23} = 1$ $y_{33} = 4$
$b_1 = 295$ $b_2 = 412$ $b_3 = 620$
$a_1 = 150$ $a_2 = 350$ $a_3 = 700$



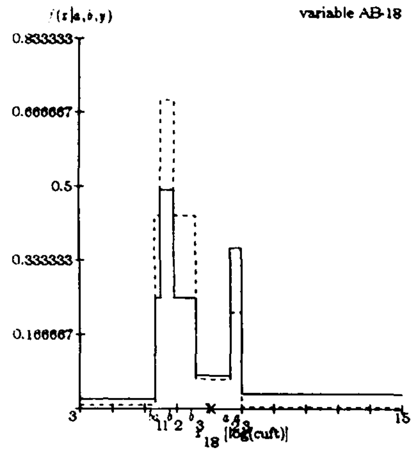
$n = 10$ true value $= 5864.25$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{21} = 0$
$y_{22} = 0$ $y_{32} = 2$ $y_{33} = 4$
$b_1 = 4000$ $b_2 = 6000$ $b_3 = 12000$
$a_1 = 2000$ $a_2 = 5000$ $a_3 = 8000$



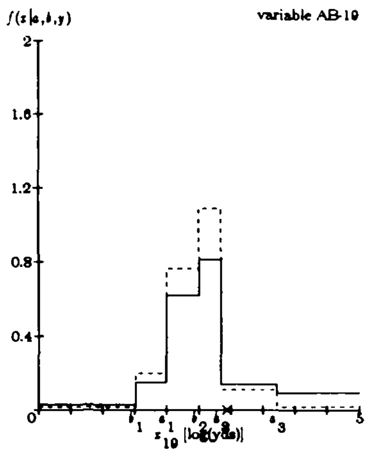
$n = 11$ true value $= 425$
$y_{00} = 1$ $y_{10} = 2$ $y_{20} = 1$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 2$ $y_{33} = 4$
$b_1 = 700$ $b_2 = 1100$ $b_3 = 1900$
$a_1 = 100$ $a_2 = 300$ $a_3 = 1000$



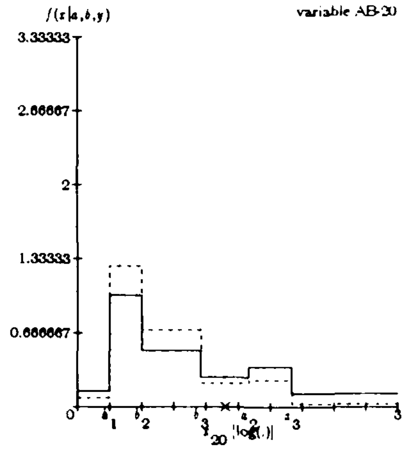
$n = 10$ true value $= 40820$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 1$ $y_{23} = 1$ $y_{33} = 4$
$b_1 = 12000$ $b_2 = 16000$ $b_3 = 22000$
$a_1 = 10000$ $a_2 = 30000$ $a_3 = 50000$



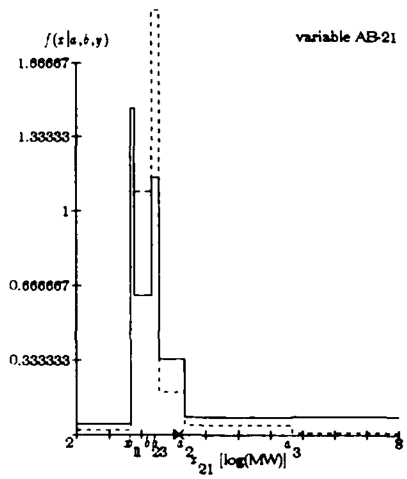
$n = 9$ true value = 7.85
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 1$ $y_{23} = 2$ $y_{33} = 4$
$b_1 = 5.8$ $b_2 = 6.5$ $b_3 = 7.3$
$a_1 = 6$ $a_2 = 8.6$ $a_3 = 9$



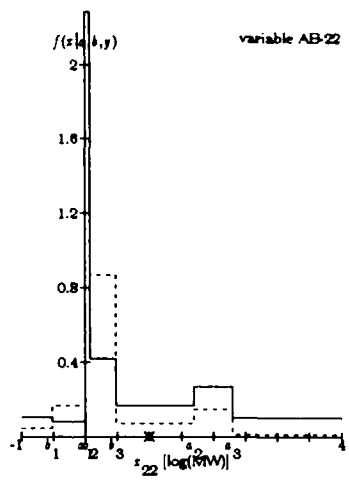
$n = 10$ true value = 2.94
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 2$ $y_{23} = 2$ $y_{33} = 4$
$b_1 = 1.52$ $b_2 = 2.5$ $b_3 = 2.85$
$a_1 = 2$ $a_2 = 2.85$ $a_3 = 3.72$



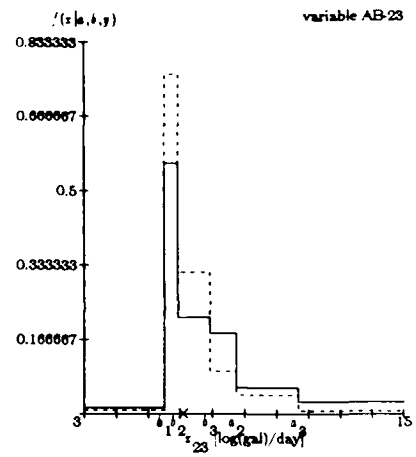
$n = 11$ true value = 1.38
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 2$ $y_{23} = 3$ $y_{33} = 4$
$b_1 = 0.3$ $b_2 = 0.6$ $b_3 = 1.15$
$a_1 = 0.3$ $a_2 = 1.6$ $a_3 = 2$



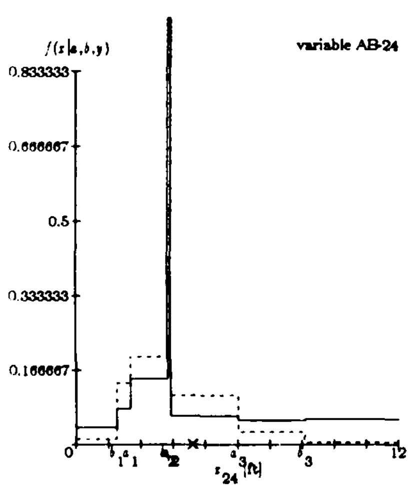
$n = 14$ true value = 3.897
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 3$ $y_{23} = 3$ $y_{33} = 4$
$b_1 = 3.08$ $b_2 = 3.4$ $b_3 = 3.54$
$a_1 = 3$ $a_2 = 4$ $a_3 = 6$



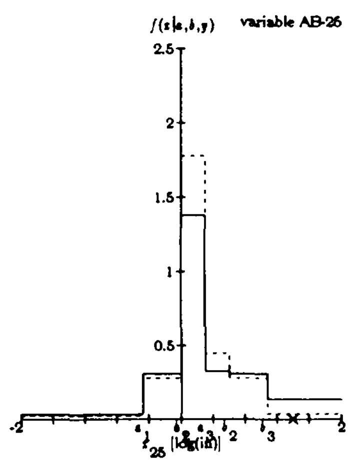
$n = 13$ true value = 1
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 4$ $y_{23} = 3$ $y_{33} = 4$
$b_1 = -0.52$ $b_2 = 0.08$ $b_3 = 0.48$
$a_1 = 0$ $a_2 = 1.7$ $a_3 = 23$



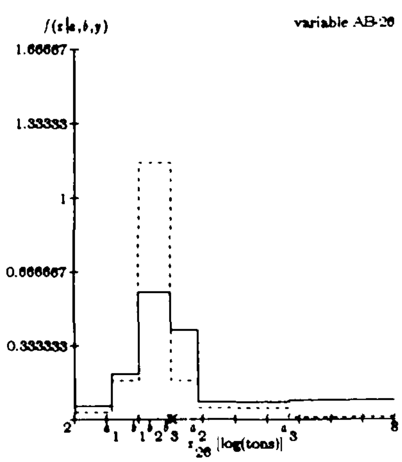
$n = 14$ true value = 6.7
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 0$
$y_{13} = 5$ $y_{23} = 3$ $y_{33} = 4$
$b_1 = 6$ $b_2 = 6.5$ $b_3 = 7.7$
$a_1 = 6$ $a_2 = 8.7$ $a_3 = 11$



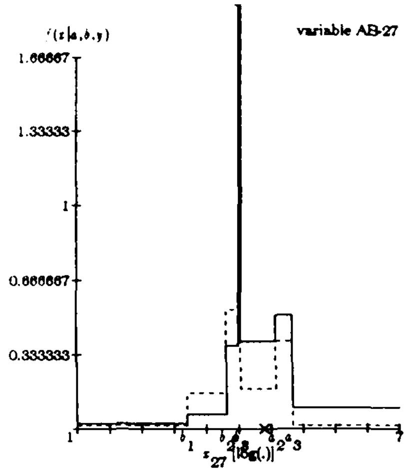
$n = 9$ true value = 4.33
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 1$
$y_{22} = 0$ $y_{32} = 2$ $y_{33} = 4$
$b_1 = 1.5$ $b_2 = 3.4$ $b_3 = 8.5$
$a_1 = 2$ $a_2 = 3.5$ $a_3 = 6$



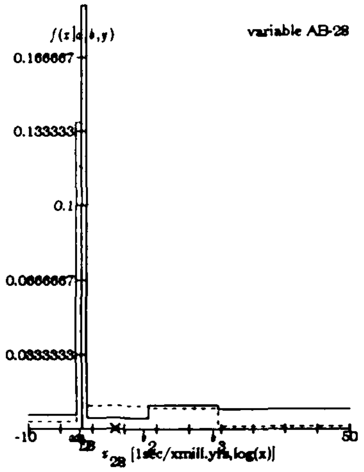
$n = 11$ true value = 1.398
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 2$ $y_{33} = 4$
$b_1 = 0$ $b_2 = 0.6$ $b_3 = 1.08$
$a_1 = -0.48$ $a_2 = 0$ $a_3 = 0.3$



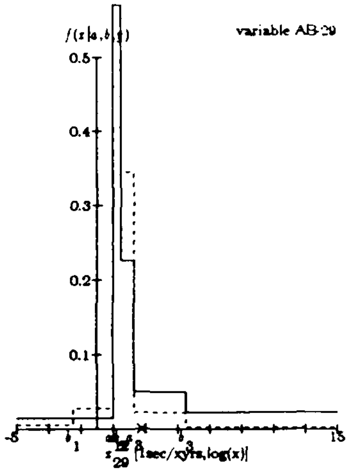
$n = 18$ true value = 3.82
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{12} = 1$
$y_{13} = 5$ $y_{23} = 3$ $y_{33} = 5$
$b_1 = 3.2$ $b_2 = 3.5$ $b_3 = 3.8$
$a_1 = 2.7$ $a_2 = 4.3$ $a_3 = 6$



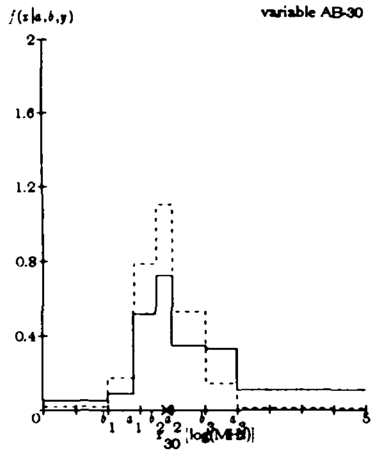
$n = 17$ true value $= 4.478$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 1$ $y_{12} = 1$
$y_{13} = 6$ $y_{23} = 3$ $y_{33} = 5$
$b_1 = 3.04$ $b_2 = 3.78$ $b_3 = 4.04$
$a_1 = 4$ $a_2 = 4.7$ $a_3 = 5$



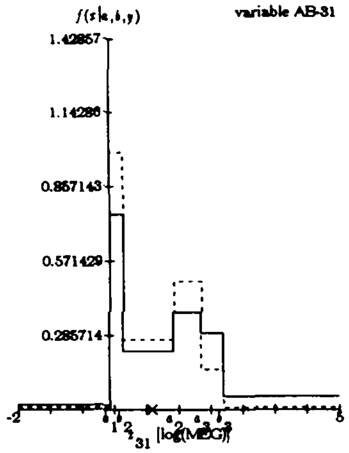
$n = 13$ true value $= 6.23$
$y_{00} = 1$ $y_{10} = 2$ $y_{20} = 2$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 2$ $y_{33} = 5$
$b_1 = 0.5$ $b_2 = 12.5$ $b_3 = 25.5$
$a_1 = -1$ $a_2 = 0$ $a_3 = 1$



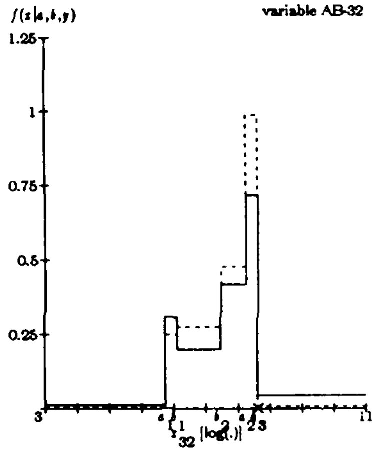
$n = 12$ true value $= 2.78$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{21} = 1$
$y_{22} = 1$ $y_{32} = 3$ $y_{33} = 5$
$b_1 = -1.5$ $b_2 = 1.5$ $b_3 = 5.5$
$a_1 = 1$ $a_2 = 1.3$ $a_3 = 2.3$



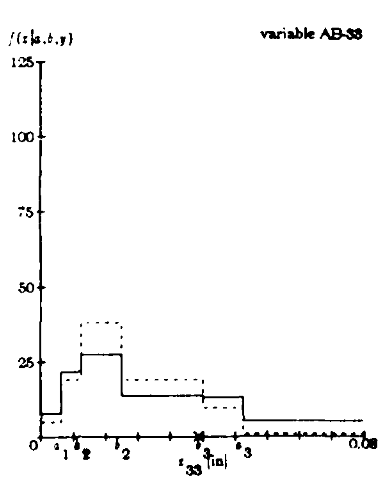
$n = 12$ true value $= 1.9$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 1$
$y_{22} = 1$ $y_{23} = 3$ $y_{33} = 5$
$b_1 = 1$ $b_2 = 1.75$ $b_3 = 2.52$
$a_1 = 1.4$ $a_2 = 2$ $a_3 = 3$



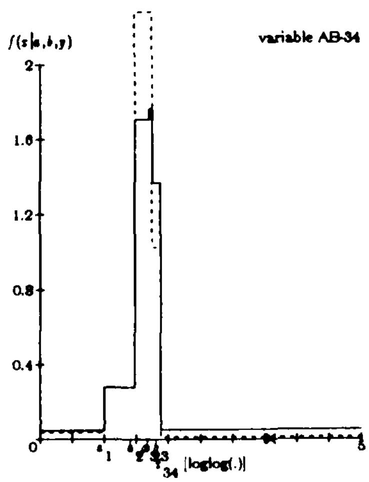
$n = 14$ true value $= 0.924$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 2$
$y_{22} = 1$ $y_{32} = 4$ $y_{33} = 5$
$b_1 = 0$ $b_2 = 0.3$ $b_3 = 2.48$
$a_1 = 0$ $a_2 = 1.4$ $a_3 = 2$



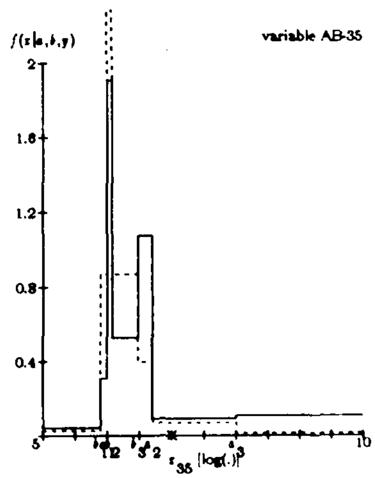
$n = 16$ true value $= 8.3$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{12} = 3$
$y_{22} = 1$ $y_{23} = 3$ $y_{33} = 5$
$b_1 = 6.3$ $b_2 = 7.38$ $b_3 = 8.3$
$a_1 = 6$ $a_2 = 8$ $a_3 = 8.3$



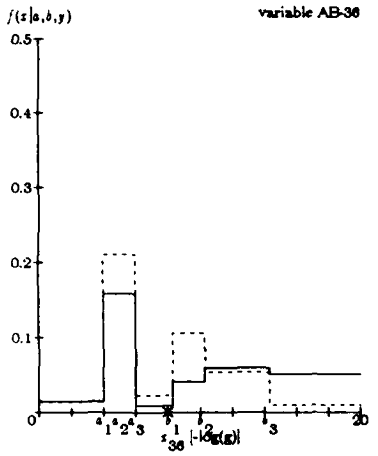
$n = 15$ true value $= 0.089$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{21} = 1$
$y_{22} = 1$ $y_{23} = 3$ $y_{33} = 6$
$b_1 = 0.01$ $b_2 = 0.02$ $b_3 = 0.04$
$a_1 = 0.005$ $a_2 = 0.01$ $a_3 = 0.05$



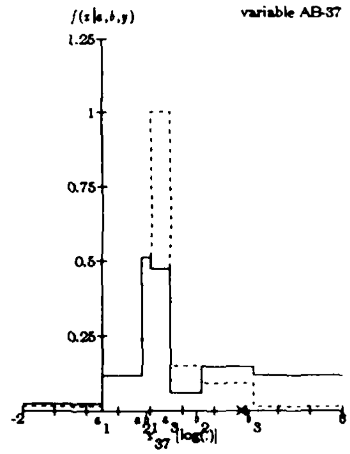
$n = 15$ true value $= 3.594$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 4$ $y_{33} = 6$
$b_1 = 1.48$ $b_2 = 1.75$ $b_3 = 1.88$
$a_1 = 1$ $a_2 = 1.48$ $a_3 = 1.7$



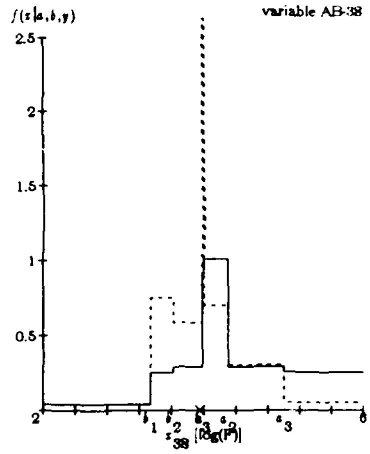
$n = 22$ true value $= 7.001$
$y_{00} = 1$ $y_{01} = 0$ $y_{11} = 1$ $y_{12} = 3$
$y_{13} = 7$ $y_{23} = 3$ $y_{33} = 7$
$b_1 = 5.9$ $b_2 = 6.08$ $b_3 = 6.48$
$a_1 = 6$ $a_2 = 6.7$ $a_3 = 8$



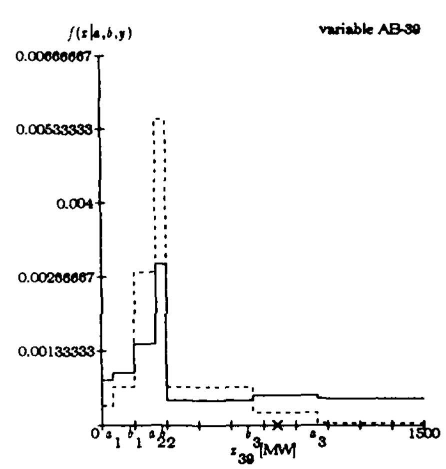
$n = 16$ true value $= 8$
$y_{00} = 1$ $y_{10} = 2$ $y_{20} = 2$ $y_{30} = 0$
$y_{31} = 0$ $y_{32} = 4$ $y_{33} = 7$
$b_1 = 8.3$ $b_2 = 10.3$ $b_3 = 14.3$
$a_1 = 4$ $a_2 = 5$ $a_3 = 6$



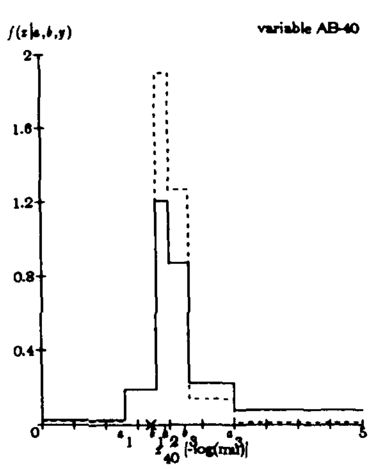
$n = 17$ true value $= 3.477$
$y_{00} = 1$ $y_{10} = 2$ $y_{20} = 2$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 4$ $y_{33} = 7$
$b_1 = 1.23$ $b_2 = 2.49$ $b_3 = 3.78$
$a_1 = 0$ $a_2 = 1$ $a_3 = 1.7$



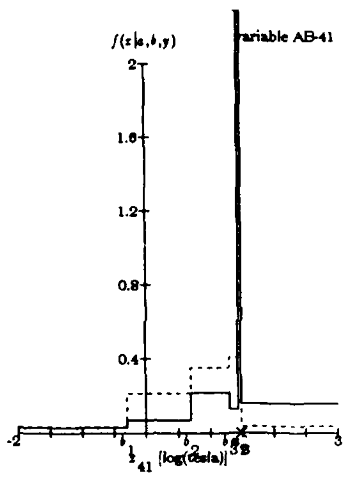
$n = 20$ true value $= 3.955$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 1$ $y_{03} = 0$
$y_{13} = 7$ $y_{23} = 4$ $y_{33} = 7$
$b_1 = 3.34$ $b_2 = 3.62$ $b_3 = 3.98$
$a_1 = 4$ $a_2 = 4.3$ $a_3 = 5$



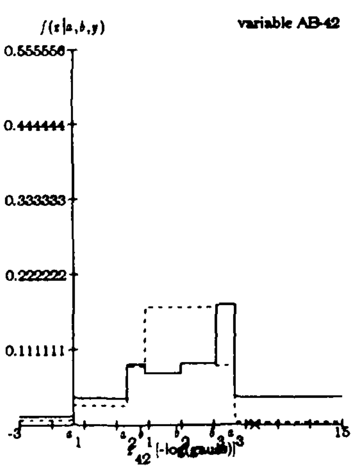
$n = 18$ true value $= 815$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{21} = 1$
$y_{22} = 2$ $y_{23} = 4$ $y_{33} = 7$
$b_1 = 150$ $b_2 = 300$ $b_3 = 700$
$a_1 = 50$ $a_2 = 250$ $a_3 = 1000$



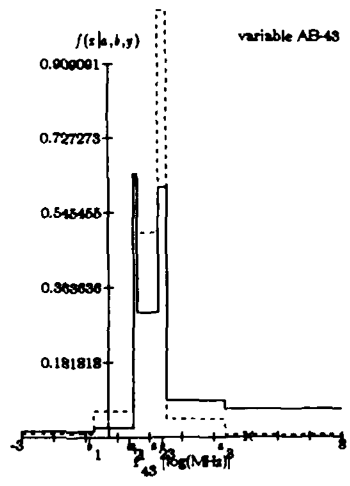
$n = 21$ true value $= 1.7$
$y_{00} = 1$ $y_{10} = 2$ $y_{11} = 1$ $y_{12} = 3$
$y_{22} = 2$ $y_{23} = 5$ $y_{33} = 7$
$b_1 = 1.8$ $b_2 = 2$ $b_3 = 2.3$
$a_1 = 1.3$ $a_2 = 2$ $a_3 = 3$



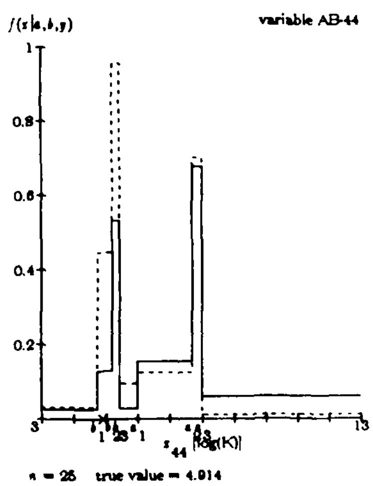
$n = 22$ true value $= 1.48$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 2$ $y_{03} = 0$
$y_{13} = 7$ $y_{23} = 5$ $y_{33} = 7$
$b_1 = -0.3$ $b_2 = 0.7$ $b_3 = 1.3$
$a_1 = 1.43$ $a_2 = 1.45$ $a_3 = 1.48$



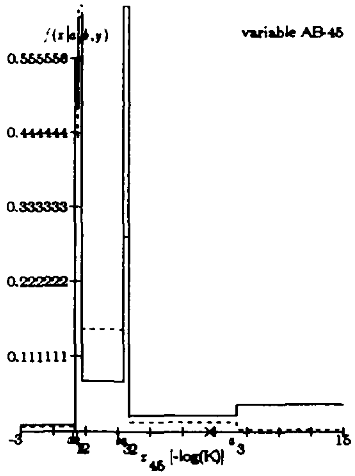
$n = 22$ true value $= 10.1$
$y_{00} = 1$ $y_{10} = 3$ $y_{20} = 2$ $y_{21} = 1$
$y_{22} = 2$ $y_{23} = 5$ $y_{33} = 8$
$b_1 = 4$ $b_2 = 6$ $b_3 = 8$
$a_1 = 0$ $a_2 = 3$ $a_3 = 9$



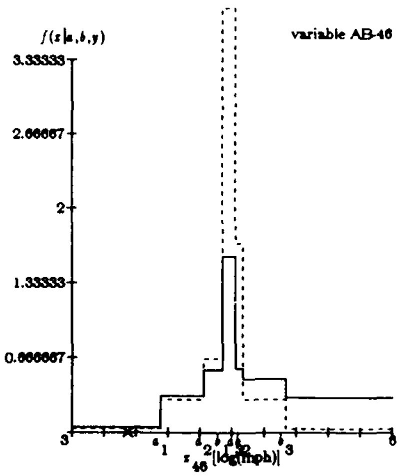
$n = 22$ true value $= 4.778$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 2$ $y_{12} = 3$
$y_{22} = 2$ $y_{23} = 5$ $y_{33} = 9$
$b_1 = -0.5$ $b_2 = 0.86$ $b_3 = 2$
$a_1 = 1$ $a_2 = 1.7$ $a_3 = 4$



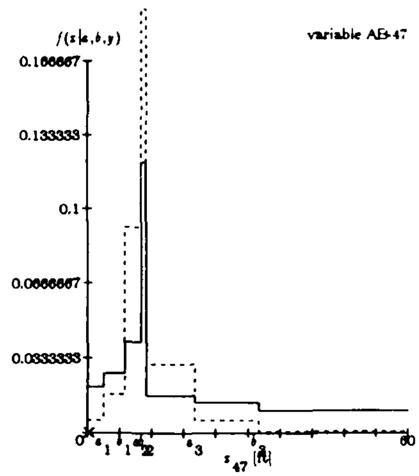
$n = 25$ true value $= 4.914$
$y_{00} = 1$ $y_{01} = 0$ $y_{02} = 2$ $y_{03} = 0$
$y_{13} = 7$ $y_{23} = 5$ $y_{33} = 10$
$b_1 = 4.75$ $b_2 = 5.22$ $b_3 = 5.44$
$a_1 = 6$ $a_2 = 7.7$ $a_3 = 8$



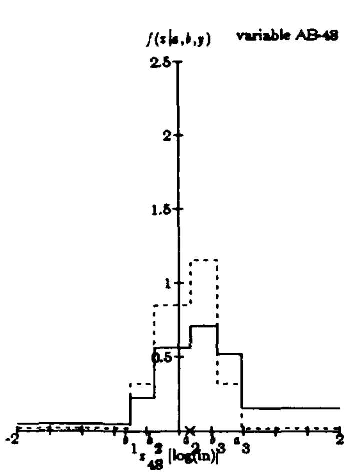
$n = 30$ true value $= 7.52$
$y_{00} = 1$ $y_{10} = 3$ $y_{11} = 1$ $y_{12} = 3$
$y_{13} = 7$ $y_{23} = 5$ $y_{33} = 10$
$b_1 = 0.2$ $b_2 = 0.4$ $b_3 = 2.7$
$a_1 = 0$ $a_2 = 3$ $a_3 = 9$



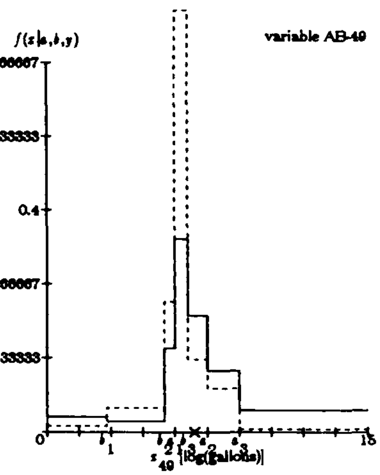
$n = 22$ true value $= 3.525$
$y_{00} = 1$ $y_{10} = 3$ $y_{20} = 2$ $y_{21} = 1$
$y_{31} = 0$ $y_{32} = 5$ $y_{33} = 10$
$b_1 = 4.41$ $b_2 = 4.6$ $b_3 = 5$
$a_1 = 3.83$ $a_2 = 4.23$ $a_3 = 4.53$



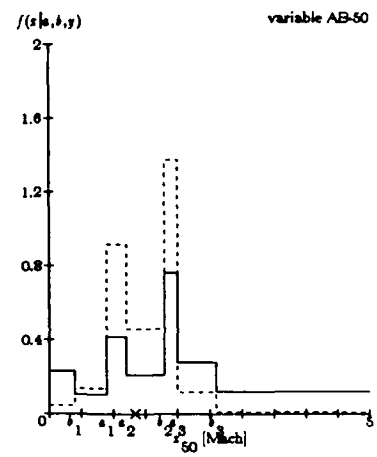
$n = 24$ true value $= 0.42$
$y_{00} = 2$ $y_{10} = 3$ $y_{11} = 1$ $y_{21} = 1$
$y_{22} = 2$ $y_{32} = 5$ $y_{33} = 10$
$b_1 = 7$ $b_2 = 11$ $b_3 = 32$
$a_1 = 3$ $a_2 = 10$ $a_3 = 20$



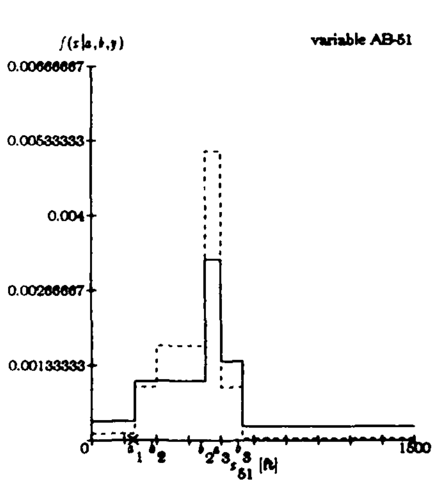
$n = 27$ true value $= 0.138$
$y_{00} = 3$ $y_{01} = 1$ $y_{02} = 2$ $y_{12} = 3$
$y_{22} = 2$ $y_{23} = 6$ $y_{33} = 10$
$b_1 = -0.6$ $b_2 = -0.3$ $b_3 = 0.48$
$a_1 = -0.3$ $a_2 = 0.15$ $a_3 = 0.78$



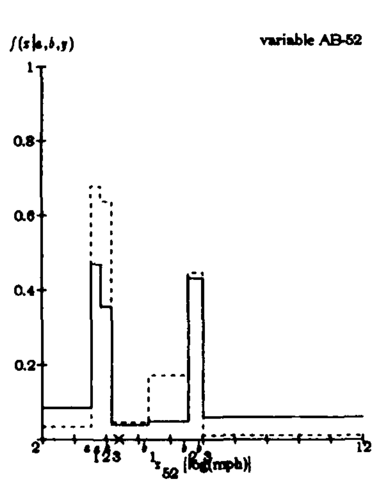
$n = 33$ true value $= 6.92$
$y_{00} = 3$ $y_{01} = 1$ $y_{02} = 2$ $y_{12} = 4$
$y_{13} = 7$ $y_{23} = 6$ $y_{33} = 10$
$b_1 = 2.8$ $b_2 = 5.5$ $b_3 = 6.6$
$a_1 = 6$ $a_2 = 7.5$ $a_3 = 9$



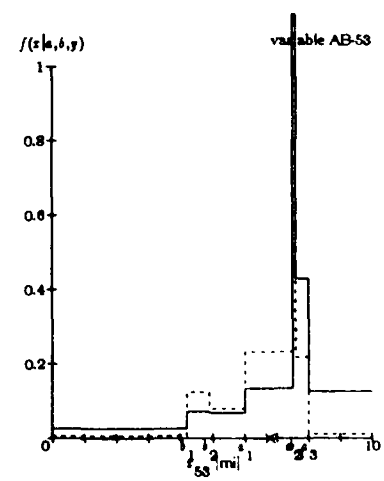
$n = 23$ true value $= 1.34$
$y_{00} = 3$ $y_{01} = 1$ $y_{11} = 1$ $y_{21} = 1$
$y_{22} = 2$ $y_{32} = 5$ $y_{33} = 10$
$b_1 = 0.4$ $b_2 = 1.8$ $b_3 = 2.6$
$a_1 = 0.9$ $a_2 = 1.2$ $a_3 = 2$



$n = 26$ true value $= 191.94$
$y_{00} = 3$ $y_{10} = 3$ $y_{11} = 1$ $y_{21} = 2$
$y_{22} = 2$ $y_{32} = 5$ $y_{33} = 10$
$b_1 = 300$ $b_2 = 526$ $b_3 = 700$
$a_1 = 200$ $a_2 = 300$ $a_3 = 600$



$n = 25$ true value $= 4.394$
$y_{00} = 4$ $y_{10} = 3$ $y_{20} = 2$ $y_{30} = 1$
$y_{31} = 0$ $y_{32} = 5$ $y_{33} = 10$
$b_1 = 5.33$ $b_2 = 6.56$ $b_3 = 7.03$
$a_1 = 3.54$ $a_2 = 3.85$ $a_3 = 4.18$



$n = 29$ true value $= 6.78$
$y_{00} = 4$ $y_{01} = 1$ $y_{02} = 2$ $y_{12} = 4$
$y_{22} = 2$ $y_{23} = 6$ $y_{33} = 10$
$b_1 = 4.2$ $b_2 = 4.9$ $b_3 = 7.6$
$a_1 = 6$ $a_2 = 7.5$ $a_3 = 8$

13. Appendix 4.

13.1 Range of the calibration indicators

Of course there are many ways to represent the calibration indicators (the z_i's) numerically. Since we are assuming that the calibration indicators are exchangeable, we will be primarily interested in the number of occurrences of a certain value in the sequence. One convenient way to obtain these counts is to choose the indicators such that the counts result when the sequence is summed.

In the single-source case the following scheme works. Let m be the number of fractiles (yielding m+1 bins). Suppose the true value fell in the jth bin. Then let the calibration indicators take on m-dimensional row-vectors such that all elements equal zero except the jth element which equals unity, or

$$\tilde{z}_i = [\underset{1}{0}, \ . \ . \ . \ , \underset{j}{1}, \ . \ . \ . \ , \underset{m}{0}]. \tag{13.1}$$

To find the counts of the hits in the bins we can define a random variable $\tilde{y}$ such that

$$\tilde{y} = \sum_{i=1}^{n} \tilde{z}_i. \tag{13.2}$$

Obviously the count of the jth bin will be in the jth position, or

$$y=[y_0, \ . \ . \ . \ , y_n] \tag{13.3}$$

In general, if we have k sources each yielding m fractiles, we can think of realizations of $\tilde{z}_i$ as k-dimensional arrays, each dimension -one for each source- consisting of an $m+1$-dimensional vector as above. With two adviser, A and B, each giving three fractiles, as in the experiment, we have a four by four matrix. One possible realization or outcome could be,

$$\tilde{z}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{13.4}$$

which would indicate that the true value fell in A's zeroth bin and B's second bin. In the experiment this was indeed the case at the second variable as can be verified e.g. in appendix 3AB.

To obtain the counts we can still apply equation (13.2), this time obtaining a k-dimensional array. The subscript, j, of the counts, y, can now be thought of as k-dimensional vectors: $j=[j_a, \ldots, j_k]$. In the experiment, j is a two-dimensional vector and the counts are of course as arranged as follows

$$\tilde{y} = \begin{bmatrix} y_{00} & y_{10} & y_{20} & y_{30} \\ y_{01} & y_{11} & y_{21} & y_{31} \\ y_{02} & y_{12} & y_{22} & y_{32} \\ y_{03} & y_{13} & y_{23} & y_{33} \end{bmatrix}. \tag{13.5}$$

13.2 Derivation of the posterior on the calibration indicator

In this section we will derive the results listed in section 7 under equations (5.1) and (5.2) using only the three form assumptions (F1, F2 and F3) and three inference assumptions (I1, I2, I3).

According to the form assumptions, we can write the posterior on the calibration indicator as follows

$$F(z_0 \mid a, \ldots, k, z_1, \ldots, z_n). \tag{13.6}$$

For the present purposes it will prove to be more convenient to work with probability mass functions, or

$$p(z_0 \mid a, \ldots, k, z_1, \ldots, z_n). \tag{13.7}$$

There are two updatings involved, one on the calibration indicators $z_1, \ldots, z_n$ and one on the information $a, \ldots, k$. We will discuss them in this order.

13.2.1 Conditioning the calibration indicators

By the definition of conditional probability we have

$$p(z_0 \mid z_1, \ldots, z_n) = \frac{p(z_0, z_1, \ldots, z_n)}{p(z_1, \ldots, z_n)}. \tag{13.8}$$

Since the z_i's (i=0, ..., n) are exchangeable (I2) the denominator of the quotient can be expanded in terms of a $(m+1)^k$-dimensional parameter array θ with

$$\sum_j \theta_j = 1. \tag{13.9}$$

De Finetti's theorem (see e.g. de Finetti [1964]) guarantees the existence of a ($(m+1)^k$-variate) distribution $F(\theta)$ such that[1]

$$p(z_0 \mid z_1, \ldots, z_n) = \frac{\int_0^1 \theta_j \cdot \prod_k \theta_k^{y_k} \cdot dF(\theta)}{p(z_1, \ldots, z_n)}, \tag{13.10}$$

where k runs over the elements of θ ($(m+1)^k$ of them), and z_0, as usual, represents the event of the true value falling in bin j. If we can now show that I1 and I3 determine the form of $f(\theta)$ we will be done in principle.

Filling in n=0 (and therefore $\tilde{y}$=0) in (13.10) we find, after some straightforward manipulation

$$p(z_0) = E(\theta). \tag{13.11}$$

However, the left side of (13.11) is simply the prior on the calibration indicator which is determined, via I1, by the claims of the sources. The first moment of $f(\theta)$ is thus determined. In general, i.e., filling $y_j = n$ in (13.10) we find that for the relation between the distribution on the calibration indicators and the moments of $f(\theta)$

1. In the following equation 1 and 0 represent vectors of 1's and 0's, respectively. In this appendix we adhere to the convention that numbers take the dimensions of the variable they are equated to. E.g. y = 0 means that y equals an array (of the dimensions of y) made entirely of zeros.

$$p(z_0, \ldots, z_n) = E(\theta^{n+1}). \tag{13.12}$$

As mentioned, I1 fixes the case for which $n=0$, leaving still an infinite number of moments to be assessed.[2] This will be done by I3. I3 assumes that the operator is as unconfident as possible about I1. We shall now have to be a little more precise about this assumption. Before we determine what as unconfident as possible means, i.e. under the constraint imposed by (13.11), we shall first fix what we mean by unconstrained unconfidence.

We shall assume that the least confident state of information is reflected by an ($(m+1)^k$-variate) uniform density over θ. This assumption is quite a bit less arbitrary than it looks at first sight. Indeed *if* we can show that (1) it is reasonable to assume that the physical property θ represents should indeed be uniformly distributed ***and*** (2) that this distribution is invariant under the admissable transformations of the scale of θ, ***then*** this assumption is meaningful.

(1) The proper interpretation of θ is that it represents the limiting frequencies of the hits in the bins, or the limiting hit rates (as $n \rightarrow \infty$). Classical statisticians prefer to call θ the fixed but unknown underlying probabilities of a hit in the various bins. From a strict Bayesian point of view this kind of wording does not make much sense and is even a bit misleading. Probabilities are never unknown (in the sense of above) since they can be derived, at least in principle, from a individual's preference among actions. Secondly, such terminology will force us to make sense of a probability of a probability (equations (13.10) and (13.11)). Even if we succeeded in proposing an adequate operational definition of this particular conceptual twister, we would also need to clarify the meaning of a probability of a probability of a probability etc. into an infinite regress.

The interpretation in terms of limiting hit rates can be made very compelling by investigating the behavior of $f(\theta \mid z_1, \ldots, z_n)$ as n grows without bound. We find after some analysis that, under mild regularity conditions[3]

$$\lim_{n \rightarrow \infty} f(\theta \mid z_1, \ldots, z_n) = \delta_{y/n}, \tag{13.13}$$

where the notation $\delta_{y/n}$ stands for the unit impulse at $\theta = y/n$. Thus, in the long run, the operator will become convinced that the parameters of the process equal the limiting hit rates. It should be borne in mind that this is only true for $n = \infty$. For bounded n the operator, in general, has a non- degenerate distribution over the parameter θ, expressing his uncertainty concerning the value of the parameter. The parameter parametrizes all possible multinomial models that might describe the multinomial process of true values falling in a set of bins. Thus, the assumption boils down to assuming that all multinomial models describing the process are equally likely.

This kind of assumption seems reasonable and is almost always made with processes like tossing coins, rolling dice and similar paradigms. As a historical precedence, it is maybe noteworthy that Laplace, in his famous calculation of the probability of the sun rising tomorrow, (implicitly) made the same assumption,[4] and so did Bayes himself in the paper that gave his name to Bayesian statistics.

(2) As for the invariance of the distribution of θ under the admissible transformation of its scale, it suffices to notice that, whether one wishes to interpret the parameter as a limiting frequency or as

2. Since $f(\theta)$ is concentrated on $[0,1]$ its moments uniquely specify its distribution.
3. That is as long as the density on theta is never zero.
4. His result was: p(sun rises tomorrow | has risen on y out of n days) = (y+1) / (n+2). Compare this with equation (13.18).

an underlying "true" probability, it is measured on an absolute scale so there are no admissible transformations in the first place. (See e.g. Pfanzagl [1968] for the measurement theoretic details.)

Accepting that the uniform over θ appropriately reflects the unconstrained least confident belief state, we will proceed to determine what the least confident belief state is under the constraint imposed by equation (13.11). Roughly speaking, what we are looking for is a distribution on the parameter that is as "close" as possible to the uniform. As usual we will measure "closeness" in distribution space with the relative information measure. Precisely speaking, we wish to find the $f(\theta)$ that minimizes

$$\int_0^1 f(\theta)\log \frac{f(\theta)}{f_{unif}(\theta)}\, d\theta \tag{13.14}$$

under the first moment constraint imposed by (13.11). The solution to this problem is easily established with the calculus of variations. One finds that, on [0,1],

$$f(\theta)=\frac{\lambda\ e^{\lambda\theta}}{1-e^{\lambda}} \tag{13.15}$$

where the values of the multiplier λ follows from the constraint as in (13.11). The product $\lambda\theta$ is shorthand for $\sum_j \lambda_j \cdot \theta_j$.

In the single-source single-fractile case, the result can be pictured, since a single θ_j (either θ_0 or θ_1) is sufficient to sepcify the distribution on θ. Now this "maximum entropy" distribution is a renormalized exponential with its tail from 1 onward clipped off. Figure 13.1 shows, in the solid lines, these densities for a number of values of the expectation. It is interesting to note that for $E(\theta)$=.5, the maximum entropy distribution equals the uniform. This is, of course, because for this value of the expectation (13.11) does not effectively constrain the minimization. Furthermore, as one would expect, this family of densities is symmetrical around the uniform. Compare e.g. the density for the values of $E(\theta)$=.7 with .3.

Theoretically speaking, the problem is now solved since we just need to plug (13.15) into (13.10) and to carry out the integration. In the single-source, single-fractile case, i.e. the cases in which z, y and θ contain two elements, the resulting integral is an integral form of Kummer's function. No closed form solution exists, although it has been extensively tabulated, a fact which is of limited use for computer implementation of the aiding system. In all other, higher dimensional cases, numerical solutions are the only alternative.

To speed up computations for computer implementation and to gain physical insight in the solution, we will search for a distribution that approximates the maximum entropy distribution closely enough for practical purposes and which will simplify the integral in (13.10). If $f(\theta)$ was a ($(m+1)^k$-variate) Dirichlet distribution with parameter r which is a by now usual $(m+1)^k$-dimensional array of r_j's. Customarily one also defines a parameter m in terms of r as follows

$$m=\sum_j r_j. \tag{13.16}$$

In other words, consider

$$f(\theta) = f_{dir}(\theta) = \frac{\Gamma(\sum_j r_j)}{\prod_j \Gamma(r_j)} \cdot \prod_j \theta_j^{r_j} \tag{13.17}$$

then the integral in (13.10) simplifies considerably. Indeed after some straightforward manipulation we find that (we also replace $z_0, \ldots, z_n$ by y) for the true value of $\tilde{x}_0$ falling in bin j.

$$p(z_0 \mid y) = \frac{y_j + r_j}{n + m} \tag{13.18}$$

I1 imposes a constraint on the parameters. Filling in $y=n=0$ in (13.18) we find

$$p(z_0) = \frac{r}{m} \tag{13.19}$$

the left side of which is determined by I1.

The fact that the integral solves elegantly using a Dirichlet distribution says of course absolutely nothing about how well this family of distributions approximates the maximum entropy distributions which follow from the assumptions. Fortunately, the family of Dirichlets is very rich and excellent approximations exist. To find a "best" fit, we use the by now familiar relative entropy metric. Of course, for values of $(m+1)^k$ higher that 2, the fits cannot be shown in two dimension. In the two-dimensional case the Dirichlets become Beta distributions. Figure 13.1 shows the Beta fits (dashed curves) to the maximum entropy distributions (normal curves), obtained by minimizing the relative entropy of the Beta distributions with respect to the maximum entropy distribution, which was done by using a numerical procedure on a computer. The figure shows both distributions for a several values of the expectation. Note that for $E(\theta) = .5$ the fit is perfect, since the uniform is a member of both families. For other values the fit is not perfect, but may, without a doubt be called good enough for our purposes.

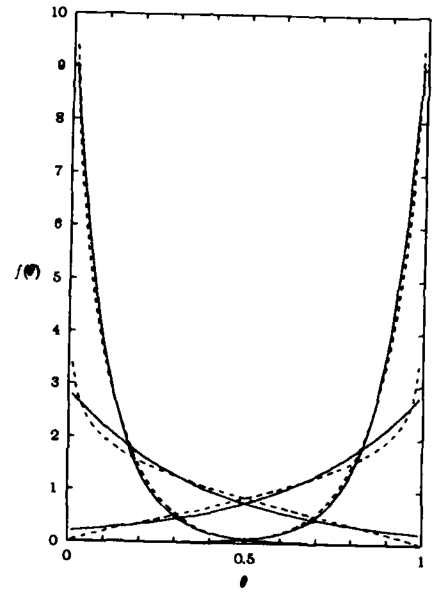


Figure 13.1. Maximum entropy densities (solid) and the approximating beta densities (dashed) for $E(\theta) = .1, .3, .5$, and $.9$.

It appears from the numerical results that for $p(z_0) \leq .5$, r is approximately 1. For $p(z_0) \geq .5$ we find that r is approximately $m-1$. (This can be also be verified by inspecting the functional form of the Dirichlet and substituting.) Filling (13.19) with this knowledge into (13.18) yields the result quoted in the document. A similar result can be obtained for $p(z_0) \geq .5$.

13.2.2 Conditioning on the information

F3 states that all the information about the sources' characteristics is contained in y. From this point of view it could be argued that $F(x_0) = F(x_0 \mid a, \ldots, k)$. A problem is that we cannot argue the same for the calibration indicators. To bear out the problem, we write

$$p(z_0 \mid a, \ldots, k) = \int_{bin\, j} dF(x_0 \mid a, \ldots, k) \tag{13.20}$$

To simplify matters a little, assume that there is an operator for whom indeed $F(x_0) = F(x_0 \mid a,\ldots,k)$. It follows that

$$p(z_0 \mid a, \ldots, k) = \int_{bin\, j} dF(x_0) \tag{13.21}$$

However, when the information has been obtained, some bins simply cannot occur anymore, so there is no reasonable $F(x_0)$ that lends positive weight to these bins. In the single-source case this only arises if the source reverses the order of its fractiles, and we could make a good case for simply rejected the information from such an incoherent source. However in the multi-source case this arises quite naturally. For instance, at variable number AB1, we have by (13.21) that

$$p(\tilde{z_0} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mid a, b) = 0. \tag{13.22}$$

The obvious solution, of course, is to use the conditioning on $a, \ldots, k$ to set the probabilities of these bins equal to zero. The naive approach that we have taken is to simply renormalize the remaining bins, or, equivalently, to assume the likelihood function to be constant on the non-zero bins. Objections can be easily raised to this solution. If a bin is very small, we would say that no $F(x_0)$ should assign much probability mass to it. However, the scales of the variables are, in general, not absolute, so a small bin can be made into an arbitrary large bin.

The proper Bayesian solution is, of course, to assume that $F(x_0)$ is known, and derive $F(z_0 \mid a, \ldots, k)$ from this distribution. This poses two practical problems. Firstly, it is very difficult to obtain $F(x_0)$ (a couple of fractiles won't do). At any rate, most operators would prefer I1 to performing this task for $\tilde{x}_0$ and all the calibration variables $\tilde{x}_1, \ldots, \tilde{x}_n$. Secondly, even if we obtained all these distributions, we would have to hope that I2 still holds. A necessary condition (not even sufficient) is that for all k, l in {0, 1,..... }, and j

$$\int_{bin\, j} F(x_k)dx_k = \int_{bin\, j} F(x_l)dx_l, \tag{13.23}$$

which is simply to good to be true. In fact, I1 relies on the operator's total unconfidence in his uncertainty (compare I3) so that he won't be to bothered by coherency constraints like (13.23).

Returning to our problem of determining the posterior on the calibration indicator, we reason the same as above, that is, $p(z_0 \mid a, \ldots, k, y)$ can then be determined from $p(z_0 \mid y)$ by setting the probabilities of the impossible bins equal to zero and renormalizing. Alternatively, we could also have conditioned on $a, \ldots, k$ first. In this case (13.19) becomes

$$p(z_0 \mid a, \ldots, k) = \frac{r'}{m'}, \tag{13.24}$$

where the primed parameters take care of the fact that certain bins have probability zero (and the rest renormalized). For the posterior and we find that

$$p(z_0 \mid a, \ldots, k, y) = \frac{y_j + r'_j}{n + m'} \tag{13.25}$$

which obviously equals the posterior obtained by conditioning in the reverse order.

13.3 Extension to degree of confidence in prior

Apart from being mathematically convenient in solving the inference problem, the Dirichlet also gives an insightful reinterpretation of the degree of confidence in the prior in terms of "equivalent calibration data". Suppose our prior for a true value falling in the zeroth bin is .5 ($=r_0/m$). We find the posterior after we have previously observed y_0 out of n falling in the zeroth bin to be

$$p(z_0 \mid a, y) = \frac{y_0 + r_0}{n+m}, \tag{13.26}$$

simply a special case of (13.18).

If we opt for being the least confident, we find that r_0=1 and m=2. Looking at the above equation, this suggests that the prior can we viewed as representing a kind of "equivalent calibration data", namely obtaining one hit in two trials. Within the constraints of the Beta distributions, these are the lowest values for the parameters that satisfy the prior constraint (i.e., that r_0/m=.5). However, there are much larger values of r_0 and m that still satisfy the prior, For instance r_0=100 and m=200. Of course, we would hardly call this a least confident prior since it will take much more "real" calibration data now to change the posterior. This suggests an operational way of solving the problem when the operator does not feal absolutely unconfident about his prior. In this case we will have to elicit an "equivalent calibration sequence" for his prior.

Of course there are other ways to elicit the necessary information. However, if we do not assume something about the form of the distribution of θ, we will have to elicit in some way or another a (countable) infinite series of moments. If an operator agrees that some family of distributions describes his confidence accurately enough, and that family has some finite number of parameters, we need only assess a finite number of moments.[5] Applications of the theory to real situations alone will show if such polishing of the theory is indeed necessary.

5. The beta has two parameters and therefore we need only asses two moments. The prior ($=r_0/m$) constitutes the first and the second, which equals $(r_0+1)/(m+1)$ the second.

14. Appendix 5

14.1 List of symbols

A	(State of information of) source A.
$A, \ldots, K$	(States of information of) sources A through K. Defying standard alphabetic enumeration we assume there are k such sources.
a	A's fractiles, i.e., $(a_1, \ldots, a_m)$ on $\tilde{x}_0$
B	(State of information of) source B.
b	B's fractiles, i.e., $(b_1, \ldots, b_m)$ on $\tilde{x}_0$.
$F(\cdot)$	The operator's marginal probability distribution function for some random variable.
$f(\cdot)$	The operator's marginal probability density function for some random variable.
F1, F2 and F3	Form assumtion 1, 2 and 3, respectively.
i	Calibration indicator or variable number, i=0,1, . . . , n.
I1, I2 and I3	Inference assumption 1, 2 and 3, respectively.
j	k-dimensional row-vector (actually, just an ordered set) $[j_a, \ldots, j_k]$.
K	(State of information of) source K.
k	Source K's fractiles.
k	Number of sources.
m	Number of fractiles elicited from each source per variable.
n	Number of calibration indicators and calibration variables.
$p(\cdot)$	The operator's probability measure.
r	$(m+1)^k$-array containing the parameters of the Dirichlet distribution.
r_j	j^{th} element of above array.
x_0	An outcome of the decision variable $\tilde{x}_0$.
$\tilde{x}_0$	Decision variable.
x_i	Outcome of calibration variable $\tilde{x}_i$, i=1,2, . . . , n.
$\tilde{x}_i$	Calibration variable, i=1,2, . . . , n.
y	Outcome of $\tilde{y}$
$\tilde{y}$	$(m+1)^k$-array containing the counts of the hits in the various bins at the n^{th} calibration indicator.
y_j	Amount of hits in bin j at the n^{th} calibration indicator.
z_0	Outcome of the calibration indicator of the decision variable.
$\tilde{z}_0$	Calibration indicator of the decision variable.
z_i	Outcome of calibration indicator $\tilde{z}_i$, i=1,2, . . . , n.

$\tilde{z}_i$	Calibration indicator, $i=1, \ldots, n$.
$\Gamma(\cdot)$	Gamma function, i.e., $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$.
θ	Outcome of $\tilde{\theta}$.
$\tilde{\theta}$	$(m+1)^k$-array containing the parameters of the multinomial model describing the stochastics of true values falling in bins.

DISTRIBUTION LIST

Engineering Psychology Program
Office of Naval Research
Code 1142EP
800 North Quincy Street
Arlington, VA 22217-5000

J. Randy Simpson
Statistics Program Code 1111SP
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Lyle D. Broemeling
Code 1111SP
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

Information Sciences Division
Code 1133
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Randal P. Schumaker
NRL A. I. Center
Code 7510 R & D Command
Naval Research Laboratory
Washington, D.C. 20375-5000.

Director
Technical Information Division
Code 2627
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D.C. 23075-5000

Dr. A. F. Norcio
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, DC 20375-5000

Human Factors Department
Code N-71
Naval Training Systems Center
Orlando, FL 32813

Mr. Norm Beck
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Engineering
Code 441
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dr. William Uttal
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. L. Chmura
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, D.C. 20375-5000

CDR C. Hutchins
Code 55
Naval Postgraduate School
Monterey, CA 93940

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 North Quincy Street
Arlington, VA 22217-5000

Professor Michael Sovereign
Joint Command, Control and
Communications Curriculum
Code 74
Naval Postgraduate School
Monterey, CA 93943

Aircrew Systems Branch
Systems Engineering Test
Directorate
U.S. Naval Test Center
Patuxent River, MD 20670

CAPT Robert Biersner
Naval Biodynamics Laboratory
Michoud Station
Box 29407
New Orleans, LA 70189

Dr. George Moeller
Human Factors Engineering Branch
Naval Submarine Base
Submarine Medical Research Lab.
Groton, CT 06340

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dr. Jerry Tobias
Auditory Research Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

LCDR T. Singer
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

LCDR R. Carter
Office of Chief on Naval Operations
(OP-01B)
Washington, D.C. 20350

CDR W. Moroney
Naval Air Development Center
Code 602
Warminster, PA 18974

Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Dr. Eugene E. Gloye
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

Dr. Robert A. Fleming
Human Factors Support Group
Naval Personnel Research
and Development Center
1411 South Fern Street
Arlington, VA 22202

Dr. Edgar M. Johnson
Tecnical Director
U.S. Army Research Institute
Alexandria, VA 22333-5600

Director, Organizations and Systems
Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Milton S. Katz
Director, Basic Research
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. A. Fregly
U.S. Air Force Office of
Scientific Research
Life Science Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332-6448

DISTRIBUTION LIST

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Dr. J. Tangney
Directorate Life Sciences
AFSOR
Bolling AFB
Washington, D.C. 20032-6448

Mr. Yale Smith
Rome Air Development
Center, RADC/COAD
Griffiss AFB
New York, NY 13441-5700

Dr. Alan Leshner
Deputy Division Director
Division of Behavioral and
Neural Sciences
National Science Foundation
1800 G. Street, N.W.
Washington, D.C. 20550

Defense Technical Information
Center (2)
Cameron Station, Bldg. 5
Alexandria, VA 22314

Dr. Amos Tversky
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Donald D. Hoffman
University of California
(Irvine)
School of Social Sciences
Irvine, CA 92717

Dr. Stanley Deutsch
NAS-National Research Council
(COFH)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Dr. Deborah Boehm-Davis
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. James T. Todd
Brandeis University
Waltham, MA 02254

Dr. David Van Essen
California Institute of Technology
Division of Biology
Pasadena, CA 91125

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D.C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Dr. Robert Wherry
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. Lola L. Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Dr. Marvin Cohen
Decision Science Consortium, Inc.
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. Scott Robertson
Catholic University
Department of Psychology
Washington, DC 20064

Dr. Andrew P. Sage
Assoc. V.P. for Academic Affairs
George Mason University
4400 University Drive
Farifax, VA 22030

Dr. James Ballas
Georgetown University
Department of Psychology
Washington, DC 20057

Dr. John Payne
Graduate School of Business
Administration
Duke University
Durham, NC 27706

Dr. Tomaso Poggio NE 43-787
MIT
Center for Biological Information
Processing
Cambridge, MA 02139

Dr. Whitman Richards
MIT E10-120
Dept. of Psychology
Cambridge, MA 02139

Dr. Robert A. Hummel
New York University
Courant Inst. of Mathematical
Sciences
251 Mercer Street
New York, NY 10012

Dr. Kent A. Stevens
University of Oregon
Dept. of Computer & Info. Sci.
Eugene, OR 97403

Dr. Donald A. Glaser
U. of California, Berkeley
Department of Molecurar Biology
Berkeley, CA 94720

END

1-87

DTIC